VOLUME 68 JUNE, 1942 NUMBER 2

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES

OF THE

UNITED STATES OF AMERICA

EXECUTIVE COMMITTEE

OF THE PROCEEDINGS

EDWIN BIDWELL WILSON, *Managing Editor*

L. A. MILLIKAN D. W. BRONK

Publication Office: Mack Printing Company, Easton, Pa.

Editorial Office: Harvard School of Public Health, Boston 15

Home Office of the Academy: Washington 25, D. C.

Entered as second-class matter, June 28, 1915, at the postoffice at Easton, Pennsylvania, under the act of August 24, 1912. Acceptance for mailing at a special rate of postage provided for in section 1103, act of October 3, 1917. Authorized June 15, 1918.

Published Monthly

Price per annum, $5.00

OFFICERS OF THE ACADEMY

PRESIDENT ..... A. A. NOYES
VICE PRESIDENT ..... L. P. EISENHART
FOREIGN SECRETARY ..... D. W. BRONK
HOME SECRETARY ..... FRED E. WRIGHT
TREASURER ..... W. J. ROSSINS

EDITORIAL BOARD

The HOME SECRETARY and the FOREIGN SECRETARY of the ACADEMY
The CHAIRMAN of the NATIONAL RESEARCH COUNCIL

A. BEVAN	A. I. HALLOWELL	W. A. NOYES, JR.
A. L. DAY	F. B. JEWETT	W. J. V. OSTERHOUT
C. A. ELVEHJEM	A. H. LAMB	F. H. SCARF
L. P. EISENHART	E. O. LAWRENCE	C. R. SOSKNSKEG
R. C. GIBBS	W. R. MILES	L. H. WEED
E. W. GOODPASTURE	J. S. NICHOLAS	S. WRIGHT

INFORMATION TO SUBSCRIBERS

Checks covering subscriptions to the PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES at the rate of $5.00 per annum should be made payable to the National Academy of Sciences, and sent to the Treasurer, National Academy of Sciences, Constitution Ave. & 21st Street, Washington 25, D. C. Single numbers $0.50.

PAST publications of the NATIONAL ACADEMY OF SCIENCES from 1863 to 1926 and of the NATIONAL RESEARCH COUNCIL from 1916 to 1925 are listed in the Index to the First Ten Volumes of these PROCEEDINGS. This Index appeared as Part 2 of Volume 13, No. 1, January, 1927, and is for sale separately at the price of $1.50. Orders should be sent to the Home Secretary, National Academy of Sciences, Constitution Ave. & 21st Street, Washington 25, D. C.

PAST volumes of the PROCEEDINGS may be obtained at five dollars per volume unless the sale of the volume would break a complete set of Volumes I to X. Single numbers may be obtained for fifty cents except where the sale of such numbers would break up a complete volume. Only two hundred complete sets of the PROCEEDINGS are available for sale—Volumes I to X, price $75.00 (including Index, *see supra*). Orders should be sent to the Home Secretary, National Academy of Sciences, Constitution Ave. & 21st Street, Washington 25, D. C.

The following publications are issued by the NATIONAL RESEARCH COUNCIL. Orders and inquiries should be addressed: Publication Office, National Research Council, Constitution Ave. & 21st Street, Washington 25, D. C.

(A) THE BULLETIN OF THE NATIONAL RESEARCH COUNCIL published at irregular intervals. Individual numbers priced variously. A discount of 20% from the list price will be given to all who subscribe for the series as issued, direct from the Publication Office.

(B) THE REPRINT AND CIRCULAR SERIES OF THE NATIONAL RESEARCH COUNCIL, individual numbers variously priced. A discount of 20% from the list price will be given to all who subscribe for the series as issued direct from the Publication Office.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 35 June 15, 1949 Number 6

VALUES AS A SUBJECT OF NATURAL SCIENCE INQUIRY

By A. L. Kroeber

Department of Anthropology, Columbia University, New York

Retired, University of California

Read before the Academy, April 25, 1949

This essay maintains the proposition that the study of values is a proper and necessary part of the study of culture, viewed as an existing part of nature.

This is said not merely as proposal or program, but as a descriptive fact holding for much of actually existing practice in anthropology and the study of culture.

Whenever a cultural fact has significance, or historical reference, it also contains a value. Significance must be distinguished from cause—from that which made a cultural phenomenon happen or come to be. Significance must also be distinguished from the end or purpose served; and from organic needs, which in their turn can be resolved either into causes or into ends of culture phenomena.

That needs—also called drives, press, imperatives and such—exist, and that they underlie and precondition culture, is indubitable. It is also obvious that culture cannot be explained or derived from needs except very partially. Hunger has to be satisfied; but *how* it is satisfied by human beings can never be derived from their being hungry, nor from their specific bodily construction. Overwhelmingly the *how* can be understood only with reference to the remainder of the culture adhered to, present and past; modified somewhat—or preconditioned—by interaction with the opportunities afforded by natural environment. Moreover, large segments of culture only begin to operate, to come into being, after the primal needs have been satisfied, have had their tensions reduced or alleviated. Such are art, religion, science. Hence these segments cannot be explained at all from physiological needs.

The essential characteristic things about a culture are its forms and patterns, the interrelations of these into an organization and the way these

parts, and the whole, work or function as a group of human beings lives under them. A culture is a way of habitual acting, feeling and thinking channeled by a society out of an infinite number and variety of potential ways of living. The particular channeling adopted is heavily preconditioned by antecedent ways and organizations or systems of culture; though it is not predetermined thereby except within certain limits. Every such system of channeling is accompanied by or contains a system of affects, which vary from place to place of their appearance, and from time to time, but some of which are usually powerful and persistent. Interconnected with these affects is a system of ideas and ideals, explicit and implicit. The combined affect-idea system of a culture at once reflects the habitual ways of action of members of the society, validates these ways to themselves and to an extent controls and modifies the ways. It is in this affect-laden idea system that, in a certain sense, the core of a culture is usually considered to reside: in it lodge its values, norms and standards—its ethos and its eidos.

When we speak of the significance of a cultural trait or item or complex of traits, what is meant is the degree to which the trait is meshed, affectively as well as structurally and functionally, to the remainder of the total system or organization that constitutes the culture. Low degree of integration normally indicates that the trait has relatively low significance for the culture as a functional unit—though it may still have considerable significance as an index of historical relationship with other cultures.

It follows that if we refuse to deal with values we are refusing to deal with what has most meaning in particular cultures as well as in human culture seen as a whole.

What we have left on elimination of values is an arid roster of cultural traits or cultural events which we are constantly tempted to animate by reintroducing the values we have banned, or else by backhandedly introducing values from our own culture. Or it is possible to attempt to explain the value-rid phenomena of the culture and their changes in terms of some causality; or possibly by a teleology.

As a matter of fact, it is and long has been prevailing practice in the description of cultures by anthropologists, or of civilizational phases by historians, to formulate the value of these cultures. Thereby the description becomes a physiognomic characterization of the culture. Such a characterization has internal import both as regards its own coherence and consistency, and external import through implicit or explicit comparison with other characterized cultures. This type of presentation, with clear-cut value designations, comprises all the most successful characterizations, or resynthesized analyses, of cultures, both by anthropologists such as Malinowski, Firth, Evans-Pritchard, and by non-professionals like Codrington on the Melanesians, Doughty in Arabia Deserta, Fustel de

Coulanges's Ancient City, Albiruni on India a thousand years ago and as far back as the Germania of Tacitus.

Reference in this matter is to values as they exist in human societies at given times and places; to values as they make their appearance in the history of our species; in short, to values as natural phenomena occurring in nature—much like the characteristic forms, qualities and abilities of animals as defined in comparative zoology. There is no reference to any absolute standard or scale of values; nor to judgments of values as better or worse—which would imply such a standard.

An absolute standard involves two qualities. First, it must be extra-natural, or supernatural, to be an *a priori* absolute. And second, ethnocentricity is implied in the elevation of any one actual standard as absolute. By contrary, standards or value-systems conceived as parts of nature are necessarily temporal and spatial, phenomenal, relative and comparative. That the first condition to the scientific study of culture is the barring of ethnocentrism has been a basic canon of anthropology for three-quarters of a century.

The forms of any culture must be described—can be appraised one might say—only in terms of their relation to the total pattern-system of that culture. The pattern system in its turn needs portrayal in terms of its total functioning and products. And so far as the pattern system is appraisable, it is in terms of comparison with the functions and results achieved by other total cultures with their respective master patterns. This is like the comparison of the total functioning and capacities of, say, an earthworm, with the functioning and capacities of other organisms.

In a sense, recognition of the functioning and capacities of an organic species in a sort of formulation of the values genetically inherent in that species. At any rate, it *can* be that, even if biologists usually are not aware of the fact and might resent the imputation of any concern with values. Further, the comparison of such values, in order to ascertain their common elements, their particularities, their apparent total range of variability, their effectivenesses and long-range permanences—such a comparison of biological values would still be within the scope of examination of natural phenomena by natural science methods.

It is as something analogous to this kind of biology or potential biology that the study of cultural forms, structures and values must be conceived. Or rather, we should say that such study has actually been made, time and again, often without explicit awareness of values being involved, and perhaps as often without awareness that the study had natural scientific significance.

It is true that values can also be viewed extra-scientifically or supernaturally. Mostly they have been so construed, with ascription to deity, soul, spirituality or a self-sufficient system of eternal, unmodifiable values

lying outside the domain of science over nature. But the present paper has no concern with such a view. Contrarily, it claims values, along with all other manifestations of culture, as being part of nature and therefore in the field of science.

A few specifications seem desirable.

There is always a gap between values and behavior, between ideals and performance. Even though values always influence behavior of cultural organisms, that is, of men, they never control it exclusively. Hence the student of culture needs to distinguish, but also to compare, ideal values and achieved behavior, as complementary to each other. The one alone falls short in substantiality, the other in significant motivation and organization of the data.

Next, values being sociocultural, they inevitably also possess psychological aspects. But as a specific quality of culture, as indeed a product of culture, their reduction to explanations in psychological terms, and of these to physiological and biochemical explanations, necessarily loses or destroys the essential specific properties of the values. These are retainable in full only as long as the phenomena of value continue to be inspected on the cultural level.

Finally, since cultural phenomena are determined in several ways—inorganically by environment, organically, psychologically and socially, as well as by existent culture—it is evident that the causality of cultural phenomena is likely to be unusually complex. Moreover they lend themselves with very great difficulty to the isolation and simplification of experiment. Within culture itself, these considerations seem to apply with even more strength to values than to say artifact production or subsistence economy. Other things equal, a descriptive or historical approach would accordingly seem more readily fruitful, in scientific inquiry into values, than any searching for causes—even immediate causes. This statement is not to be construed as a methodological ban against the study of causality in values, but as an intimation that the causal approach is inherently difficult and that valid, non-speculative results bid fair to be thin and slow. While a formal approach is thus indicated as more fruitful, this need by no means be limited to enumerative description, nor to enumeration of sequences. Beyond these, the comparison of organization, functioning and interrelations of cultural values and value systems invite methodical scientific research.

*INDIVIDUAL METABOLIC PATTERNS, ALCOHOLISM, GENETOTROPHIC DISEASES**

BY ROGER J. WILLIAMS, L. JOE BERRY AND ERNEST BEERSTECHER, JR.

THE BIOCHEMICAL INSTITUTE AND THE DEPARTMENT OF CHEMISTRY OF THE UNIVERSITY OF TEXAS AND THE CLAYTON FOUNDATION FOR RESEARCH, AUSTIN

Read before the Academy, April 27, 1949

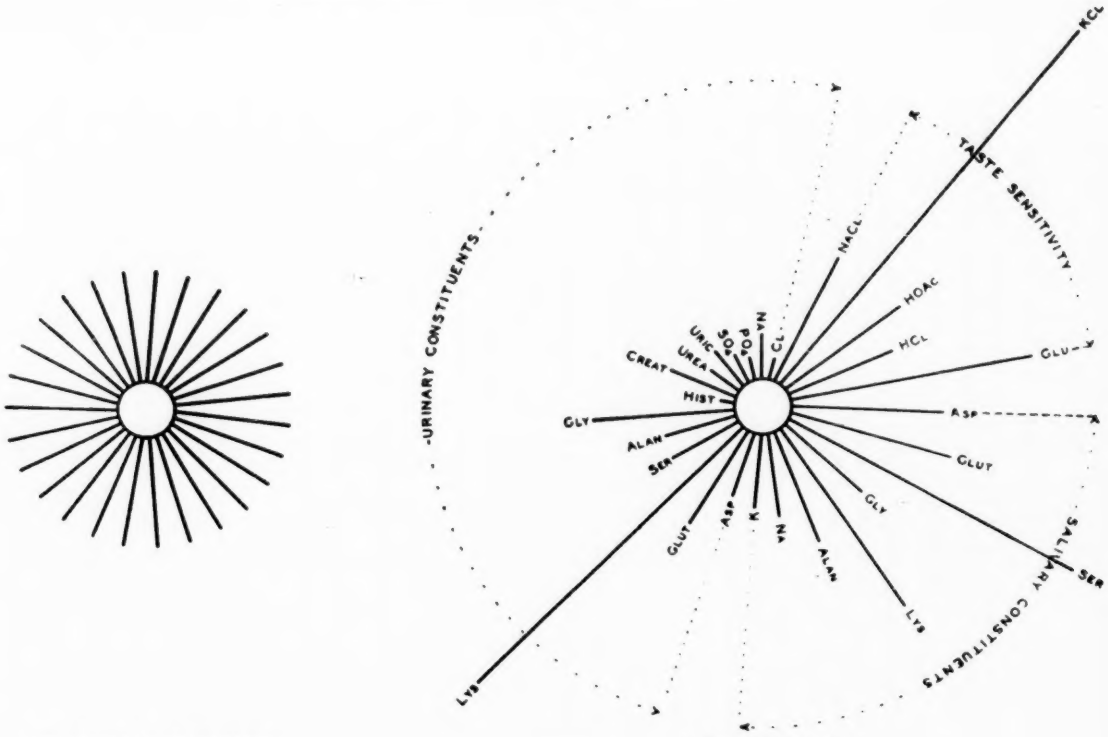
At the 1948 Annual Meeting of the National Academy there was presented preliminary evidence indicating that each individual possesses what may be called a "metabolic personality"—that is, a distinctive pattern of metabolic traits.[1] The existence of these distinctive patterns is established by analysis of body fluids and of physiological responses to chemical stimuli and is exemplified in figure 1. At the left is graphed, using polar coördinates, various metabolic traits of the "hypothetical average individual" whose every trait is the average of his fellows. At the right is graphed on the same scales the metabolic pattern of a typical real individual in which many traits vary widely from the average. This pattern is a relatively consistent one and is maintained for this individual at least over a period of several months.

The hypothesis was set forth that the metabolic patterns of individuals are of great import in the lives of the individuals who possess them, not only with respect to susceptibility to disease, but also with respect to nutrition, sex, mental abilities and many other facets of their lives.

Today we wish to present further and more specific evidence with respect to how these metabolic patterns are of moment in connection with the specific problem of alcoholism. This evidence is based primarily upon animal experimentation, the background of which is described elsewhere.[2-4]

It should be noted at the outset that individual animals in an ordinary laboratory colony have distinctive metabolic traits just as do the members of a human population. In order to conserve space we will not present data to substantiate this statement. When such animals are placed in individual cages and are continuously given a choice between water and 10% alcohol (the positions of the two drinking bottles being switched daily), they exhibit individual patterns with respect to alcohol consumption. Some drink fairly heavily beginning perhaps the first day the alcohol is offered. At the other extreme, some continue to abstain for an indefinite period of time. Intermediate between these are those that drink very moderately for an indefinite period of time and those that drink very little at first but after a few weeks drink relatively heavily. Some individual animals exhibit a relatively steady consumption day after day; in others the consumption fluctuates widely. (Figure 2)

These individual patterns of alcohol consumption are genetically con-



Hypothetical "Average" Individual

A Typical Individual

FIGURE 1

The following abbreviations are used in Figure 1: Creat = creatinine, Uric = uric acid, Hist = histidine, Gly = glycine, Alan = alanine, Ser = serine, Lys = lysine, Asp = aspartic acid, Glut = glutamic acid, Glu = glucose.

trolled as is evidenced by the fact that each strain of animals appears distinctive with respect to its alcohol consumption records. We have tested in a preliminary way three strains of mice and two strains of rats. Possibly more important evidence is the fact that when a closely inbred strain of mice resulting from the continuous use of brother-sister matings was tested, all the animals exhibited substantially the same pattern of

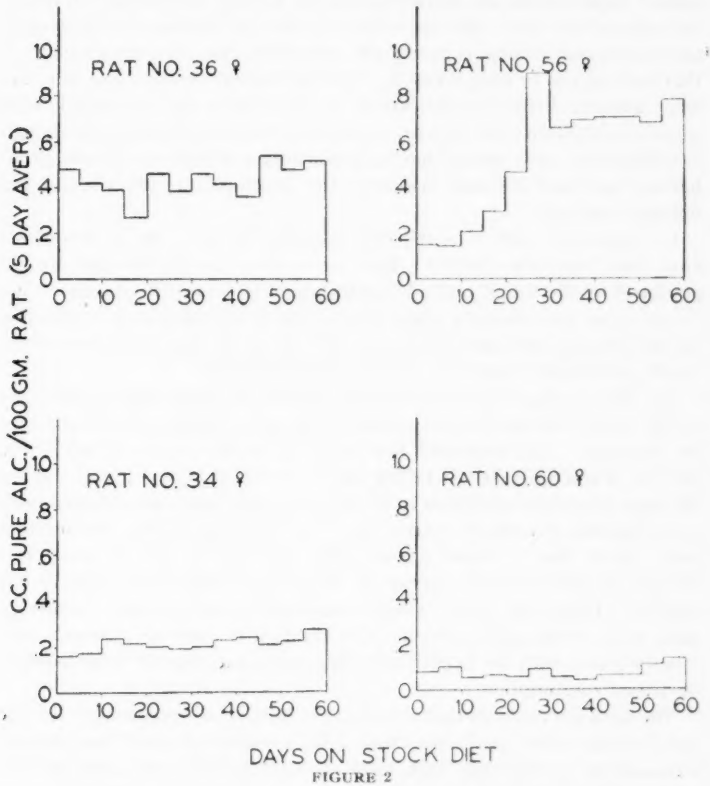


FIGURE 2

response. Further convincing evidence on this point was obtained when the different strains of animals were tested while on different diets.

There has been a far-reaching and, we believe, disastrous tendency not only on the part of social students and educators in general but also to a degree among the medical profession to avoid the possible importance of heredity as a factor in any major human problem. The attitude can be

expressed something like this: "If a certain disease is hereditary, or if mental traits and abilities are hereditary, nothing can be done about it. Therefore, let us assume that diseases are rarely hereditary and that every normal person is capable of unlimited development in almost any direction."

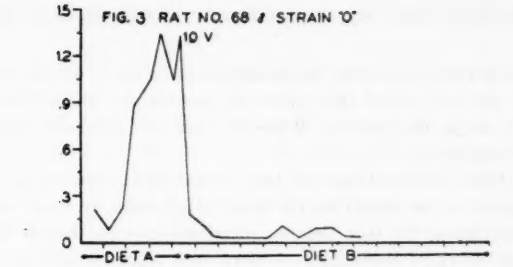
How wrong this attitude is will be amply illustrated if the results of our animal experiments are substantiated by human experience. Whereas our experiments have fully convinced us that the tendency of individual rats to consume alcohol is genetically controlled, this does not mean at all that nothing can be done about it. On the contrary, on the basis of somewhat unusual dietary considerations we have been able to abolish with surprising regularity the appetite of rats and mice for alcohol, even though it is clear that each animal has its genetically controlled metabolic peculiarities and that different factors in the treatment are effective for the different animals.

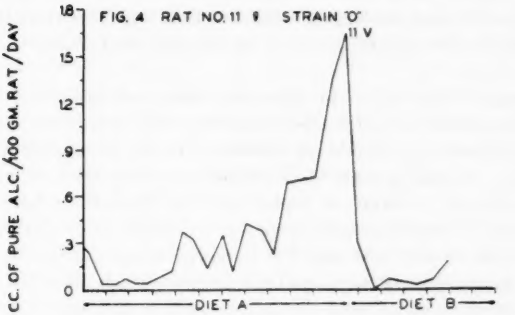
In Figures 3, 4 and 5 are depicted examples of the scores of rats which have been similarly affected. Rat No. 68 had his appetite for alcohol abolished by the simple oral administration of 10 known B vitamins. We do not know how many or which ones of the B vitamins were responsible for the effect in this particular rat. We do know, however, that only a small percentage of rats will yield to this treatment.

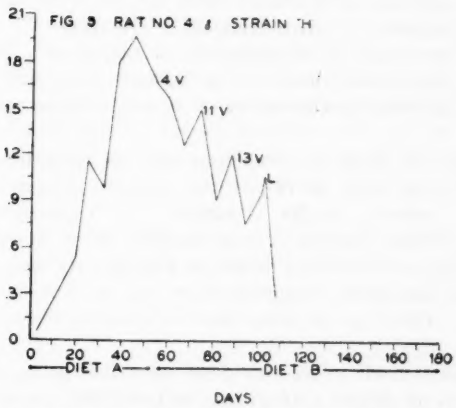
Rat No. 11, Figure 4, is one which responded, however, when in addition to the 10 B vitamins the antipernicious anemia vitamin was administered by injection. This treatment was found to be effective for 17 out of the 24 rats of our original laboratory strain, to which it was given. When the same treatment was given to 45 rats of a strain which we obtained from a neighboring laboratory, only 3, or 7%, responded. The consumption curve of rat No. 4, Figure 5, exemplifies the finding that for some rats vitamin A deficiency is a factor in the development of an appetite for alcohol. Thirty-one rats, whose consumption curves were relatively high, were administered vitamin A in addition to other supplements previously given, with the result that every single one showed immediately a decreased consumption.

The same curve shows that something in linseed oil—presumably linoleic and linolenic acids—is also a factor. Like a number of other rats, this one remained at a relatively high level of consumption even after the administration of 13 vitamins. When some linseed oil was included in the diet the consumption promptly fell to a low value.

Our experimental studies lead us to conclude that we can by dietary control keep all the rats in a colony from consuming alcohol when it is furnished *ad libitum* under our conditions. Even an initial experiment in this direction, made before we were acquainted with the effects of the antipernicious anemia vitamin, vitamin A or unsaturated fat acids, caused







Diet A is deficient in B vitamins. Diet B is the same except for added B vitamins. *4V* = thiamine, riboflavin, calcium pantothenate and folic acid (Folvite). *10V* = same as *4V* with addition of pyridoxine, choline, inositol, *p*-aminobenzoic acid, nicotinic acid and biotin. *11V* = *10V* with the addition of antipernicious anemia liver preparation given intraperitoneally. *13V* = *11V* with the addition of vitamins A and D. *L* = *13V* with the addition of linseed oil.

in a group of rats a decrease in consumption down to $^1/_7$ of the original.

We are also convinced that after an appetite for alcohol has been developed, it can be abolished in all rats by a use and extension of the method we have employed.

Out of these observations we have developed a theory of alcoholism which appears to be sound on the basis of all tests we have been able to apply. According to this theory, alcoholic craving which develops in compulsive drinkers constitutes a perverted appetite which arises as a result of one or more dietary deficiencies. Little is known as to how special appetites and cravings develop, but that they arise from the internal physiological environment has been amply demonstrated in a number of cases.

The dietary deficiencies in alcoholics arise not primarily because of failure to eat what is regarded as satisfactory food, but because the genetic patterns of these individuals are conducive to the development of crucial deficiencies. Partial genetic blocks which have recently been investigated at the California Institute of Technology[5] are thought to be responsible. When there is a partial genetic block, the capability of producing a specific enzyme is not entirely lost, but it is impaired to such an extent that there develops an augmented nutritional requirement which may be for a specific vitamin, amino acid, mineral element or other metabolic substance.

It is hoped that for the treatment of alcoholism in human beings a "shotgun" therapy may be developed which will effectively curb or abolish the excessive appetites of many compulsive drinkers.

Before treatment of alcoholics can be placed on a satisfactory basis, however, much more needs to be learned about individual metabolic patterns and what specific deficiencies are most often associated with the disease.

Out of this theory of alcoholism and the laboratory findings which support it we have developed the concept of genetotrophic disease (geneto = genetic; trophic = nutritional). A genetotrophic disease (of which we believe alcoholism is an example) is one arising fundamentally from nutritional deficiency which in turn has its basis in a genetically controlled augmented requirement for one or more specific nutritional elements. Other factors may enter to complicate or aggravate such a disease.

It suggests itself to us with great force that probably many diseases which have an obscure etiology may be essentially genetotrophic in origin. Among the diseases in which genetotrophic factors are probably involved in an important way are allergies, mental diseases, cardiovascular diseases, arthritis, multiple sclerosis and even cancer. This is hypothesis, of course, but it is based upon a large body of data.

Let us look for a few minutes at the tremendous problem of mental

disease which has received wide-spread and well-deserved popular attention in recent years.

If we consider an example of a specific, important type of mental disease—schizophrenia—we are confronted by irresistable evidence that it has physiological concomitants. If we take a one-sided view, we will assume *a priori* that the physiological disturbances are psychogenic in origin and are a result of mental derangement. If we look at the problem more broadly, however, we will consider the possibility that mental derangement may be the result of physiological impairment. We believe there is abundant reason for investigating this latter possibility.

We have made a preliminary study of the metabolic patterns of a considerable number of patients of the Austin State Hospital. This study has revealed the existence among these patients of a number of metabolic traits which are either absent or rare in the well population. We are of the opinion that some of these traits will be found to be important in connection with the causation of mental diseases, and that such investigations need to be greatly extended.

The possibility that mental diseases are genetotrophic in origin or that they have genetotrophic factors in their etiology cannot be brushed aside lightly in view of the known fact that certain mental difficulties have been treated with remarkable success by physiological means. I refer to the severe mental difficulties that sometimes accompany menopause which can be completely dispelled by hormone therapy, and the mental derangement accompanying pellagra which can be cured by dietary means.

The recent survey by Kallmann[6] gives what appears to be practically irrefutable evidence that schizophrenia has a genetic basis. This by no means indicates that nothing can be done about it. When and if we find that certain metabolic difficulties are at the basis of the trouble, there is, we believe, an excellent chance that a way out can be found. If genetotrophic factors are found, as seems highly probable, certainly help can be given.

In any event the physiological approach to the study of mental disease needs far more emphasis than it has so far received. We cannot afford to turn the problem over to cultists or to those who are wholly unprepared for physiological exploration.

* This research was supported in part by grants from the Research Corporation and the Research Council on Problems of Alcohol, New York.

[1] Williams, R. J., *Science*, **107**, 459 (1948).

[2] Williams, R. J., Berry, L. J., and Beerstecher, E., Jr., *Arch. Biochem.* (in press).

[3] Mardones, J., Segovia, N., and Onfray, E., *Arch. Biochem.*, **9**, 401 (1946).

[4] Williams, R. J., *Chem. Eng. News*, **25**, 1112–1123 (1947).

[5] Mitchell, H. K., and Houlahan, M. B., *Fed. Proc.*, **6**, 506–509 (1947).

[6] Kallmann, F. J., "Personality, Society and Culture," (Kluckholn, C., and Murray, H. A., editors), A. H. Knopf, New York, N. Y., 1948, pp. 60–79.

INHIBITION OF PLANT GROWTH BY PROTOANEMONIN AND COUMARIN, AND ITS PREVENTION BY BAL*

BY KENNETH V. THIMANN AND WALTER D. BONNER, JR.

BIOLOGICAL LABORATORIES, HARVARD UNIVERSITY, CAMBRIDGE, MASS.

Read before the Academy, April 27, 1945

1. Introduction.—In two papers dealing with the growth by cell enlargement of sections of oat coleoptiles and pea stems, growing in simple solutions in the dark, the effects of several inhibitors were studied.[1, 2] It was deduced that the elongation of these isolated plant parts is mediated through a sulfhydryl enzyme, which is particularly open to attack by such substances as arsenite, iodoacetate and organic mercurials. These reagents are powerful growth inhibitors. In the case of iodoacetate the addition of malate or other organic acids can protect against the inhibitor, and there is other evidence that the sensitive enzyme is connected with organic acid metabolism. It also appears that this enzyme, or the reaction it catalyzes, is the main limiting factor in growth of these plants. Furthermore, it was tentatively suggested that the unsaturated lactones, which are well known to inhibit growth of seedlings, also do so through the same sulfhydryl enzyme. Although reasons were given for this view, no evidence was offered at the time. The present paper gives experimental evidence which, we believe, confirms the suggestion.

In the meantime a study of the literature has revealed evidence already in existence for an interaction between unsaturated lactones and thiol groups. Hauschka, Toennies and Swain[3] found that cysteine antagonized the action of δ-hexenolactone on the flat-worm *Dugesia tigrina*, and chemical evidence for a reversible reaction between the same lactone and cysteine has been given by Cavallito and Haskell.[4] Geiger and Conn[5] have shown that clavacin (patulin), an antibiotic of unsaturated lactone structure, is inactivated by cysteine and thioglycollate. This evidence strengthens the view proposed above, and points in the same direction as our own experiments on higher plants.

2. The Inhibiting Action of Unsaturated Lactones on Growth.—The inhibiting effect of a number of unsaturated lactones on seed germination and root elongation has been reported by many workers (see Audus and Quastel[6] and literature there cited). However, seed germination is an obscure process and appears to involve root elongation, which is auxin-sensitive, as well as the complex of enzyme activations participating in germination proper. Fortunately the lactones also exert a clear-cut inhibition on a well-studied growth process. This was first made clear by Veldstra and Havinga[7] using coumarin as the unsaturated lactone, and

studying the curvature of slit pea stems in auxin solutions. We have confirmed the effect on slit pea stems on several occasions (cf. reference 8); curvature is inhibited by concentrations of 1.10^{-4} *M* and above; the inhibition is complete, *i.e.*, the inward curvature is abolished, at about 3.10^{-3} *M*. The procedure of these experiments was as previously described.[9] As with arsenite (cf. reference 2, Fig. 5) the threshold concentration depends somewhat on the concentration of auxin used; *i.e.*, the higher the auxin concentration the larger the curvature of the uninhibited controls and the lower the concentration of coumarin which shows a detectable effect. The inhibition is not removed by adding malate or other organic acids.

Protoanemonin, the lactone of 3-hydroxy-Δ-^{1,3}-pentadienoic acid, which was kindly made available to us by Dr. Chester Stock, behaved much like coumarin but was effective at 10 to 30 times lower concentra-

TABLE 1

INHIBITION OF SLIT PEA STEM CURVATURES BY PROTOANEMONIN. THE CURVATURES ARE MEASURED BY THE "STEM-REFERENCE" METHOD;[9] INWARD CURVATURE IS DESIGNATED AS POSITIVE, OUTWARD AS NEGATIVE

CONCENTRATION OF INDOLE-ACETIC, MG./L.	CONCENTRATION OF PROTOANEMONIN, $\times 10^{-4}$ M	CURVATURES, DEGREES
1	0	+112
1	0.04	+ 87
1	0.2	+ 51
1	1.0	−118
1	4.5	Toxic
0	0	−116

tions; inhibition was detectable at 4.10^{-6} *M* (0.4 p. p. m.) and complete at 1.10^{-4} *M*† (see table 1).

Both lactones also inhibit growth of coleoptile sections, and do this at concentrations lower than those effective in the pea test. The technique of these experiments, using 3-mm. sections cut from 72-hour-old etiolated Avena coleoptiles, has been described in detail elsewhere.[1] Figure 1 summarizes the data of three experiments, one on coumarin and two on protoanemonin, using indole-acetic, 1 mg. per liter, and sucrose, 1%, as growth medium. To make the experiments comparable the growth of the control sections in the absence of lactone is set at 100%. The data show that low concentrations of both lactones accelerate growth, protoanemonin being more effective than coumarin. They also reveal that when inhibition begins to occur the remaining growth is linearly proportional to the logarithm of the concentration of lactone. The lines are so evidently straight that there can be no doubt of the relationship. The slope is different for the two lactones.

The shapes of these curves, with their marked growth promotion at

low concentrations, are similar to those previously recorded for iodoacetate.[1] However, the inhibition is not prevented by malate or succinate. The effects of the lactones on growth are thus intermediate between those of iodoacetate and of arsenite.

It may be mentioned that Penicillin G, which has been thought also to react with SH groups, did not inhibit growth in concentrations up to 10^{-4} *M*.

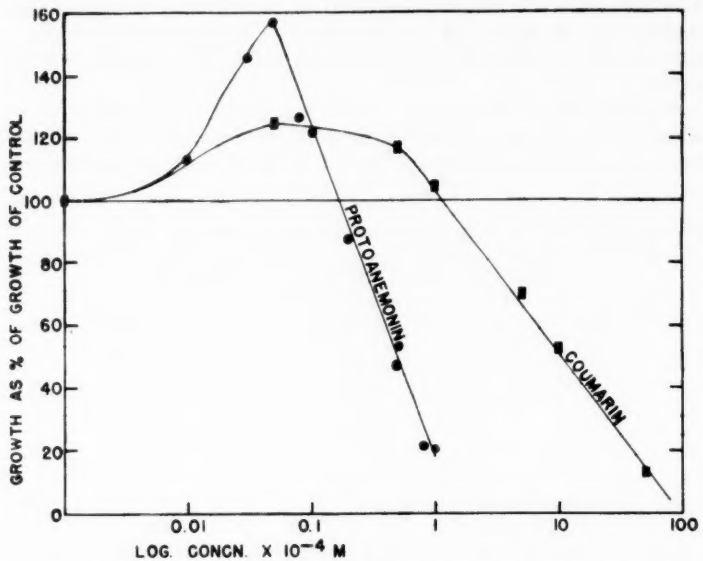


FIGURE 1

The effect of coumarin and protoanemonin on the growth of 3-mm. Avena coleoptile sections. The total growth after 48 hours, as per cent of the control growth, is plotted as a function of the log of the lactone concentration. Sections cut from 72-hour coleoptiles, breaking surface. All solutions contained sucrose, 1%, and 1 mg./l. indoleacetic acid.

3. Protection Against the Inhibition.—When the sections are treated with either one of the lactones and at the same time with 1,2-dimercaptopropane (BAL) the inhibition is very largely prevented. It has not been possible to achieve complete protection because the BAL itself has a toxic action above 3.10^{-4} *M*, but the effects are nevertheless quite clear-cut.

Table 2 shows an experiment with pea curvatures using protoanemonin at a concentration sufficient to give an inhibition of well over 50%. The BAL at 3.10^{-4} *M*, which by itself reduces the curvature 30°, is able to bring up the curvature in presence of protoanemonin from −39° to +121°.

Butanedithiol had a similar effect but the results were more erratic.

Table 3 shows a similar experiment using coumarin. The increases in curvature due to BAL are again very large, and in the first case bring the value above that of the controls. In this table the extent of the protection by a given concentration of BAL is more or less independent of the coumarin concentration.

With the growth of coleoptile sections, protection against coumarin was also real, but quantitatively less impressive. Coumarin at 10^{-3} M

TABLE 2

INHIBITION OF CURVATURES OF SLIT PEA STEMS BY PROTOANEMONIN AND ITS PREVENTION BY BAL

CONCENTRATION OF INDOLE-ACETIC, MG./L.	CONCENTRATION OF PROTOANEMONIN, $\times 10^{-4}$ M	CONCENTRATION OF BAL, $\times 10^{-4}$ M	CURVATURE, DEGREES	INCREASE DUE TO BAL, DEGREES
10	0	0	+238	...
10	0.4	0	− 39	...
10	0.4	0.5	+ 73	+112
10	0.4	3	+121	+160
10	0	0.5	+271	...
10	0	3	+207	...
0	0	0	−183	...

TABLE 3

INHIBITION OF CURVATURE OF SLIT PEA STEMS BY COUMARIN, AND ITS PREVENTION BY BAL

CONCENTRATION OF INDOLE-ACETIC, MG./L.	CONCENTRATION OF COUMARIN, $\times 10^{-4}$ M	CONCENTRATION OF BAL, $\times 10^{-4}$ M	CURVATURE, DEGREES	INCREASE DUE TO BAL, DEGREES
10	0	0	+234	...
10	1	0	+165	...
10	1	3	+290	+125
10	3	0	− 43	...
10	3	3	+ 84	+127
10	10	0	− 65	...
10	10	3	− 34	+ 31
10	30	0	−172	...
10	30	3	+ 2	+174
0	0	0	−199	...

reduced the growth to 49% (cf. 53% in Fig. 1) and addition of BAL 1.10^{-4} M raised this to 61%. Table 4, however, shows excellent protection of coleoptile sections against the inhibition due to protoanemonin. It will be seen that growth which has been reduced to one-third that of the controls is increased to 88% of the control by BAL at 1.10^{-4} molal.

When the inhibition is more nearly complete the protection against it by BAL or butanedithiol is quantitatively smaller. With peas in 10^{-4} M protoanemonin the effect of BAL was practically zero, and a similar

reduction of effect is seen in the last two lines of table 4, in which, however, the protection is still definitely present.

4. Summary and Conclusion.—A study is reported of the action of coumarin and protoanemonin on the growth of two standard objects—sections of Avena coleoptiles elongating in auxin and sucrose, and slit internodes of Pisum stems curving in auxin alone. Both lactones inhibit these growth processes, the protoanemonin being 10 to 30 times the more

TABLE 4

INHIBITION OF GROWTH OF 3-MM. COLEOPTILE SECTIONS BY PROTOANEMONIN AND ITS PREVENTION BY BAL. ALL SOLUTIONS CONTAIN INDOLE-ACETIC ACID 1 MG./L. AND SUCROSE 1%

CONCENTRATION OF PROTOANEMONIN, $\times 10^{-4}$ M	CONCENTRATION OF BAL, $\times 10^{-4}$ M	GROWTH AS PER CENT INCREASE OVER THE INITIAL LENGTH	GROWTH AS PER CENT OF GROWTH OF CONTROLS	INCREASE DUE TO BAL, %
0	0	91.0	(100)	..
0.5	0	30.1	33	..
0.5	1	79.9	88	+55
0.5	3	71.3	78	+45
1.0	0	14.6	16	..
1.0	3	31.8	35	+19

active. At subinhibiting concentrations the growth is promoted in each case. The inhibition is clearly prevented by the addition of dimercaptopropanol, BAL, at concentrations below 3.10^{-4} molal.

It is concluded that coumarin and protoanemonin inhibit growth through reacting with a sulfhydryl enzyme, and that this enzyme is normally a limiting factor in growth, and is very probably the same as that inhibited by iodoacetate and arsenite.

* This work was made possible through a grant from the Committee on Growth of the National Research Council, acting for the American Cancer Society.

† The instability of the lactone in aqueous solution, even at 1°C., limited the number of possible experiments.

[1] Thimann, K. V., and Bonner, W. D., I, *Am. J. Bot.*, **35**, No. 5, 271–281 (1948).
[2] Thimann, K. V., and Bonner, W. D., II, *Ibid.*, **36**, No. 2, 214 (1949).
[3] Hauschka, T., Toennies, G., and Swain, A. P., *Science*, **101**, 383 (1945).
[4] Cavallito, C. J., and Haskell, T. H., *J. Am. Chem. Soc.*, **67**, 1991 (1945).
[5] Geiger, W. B., and Conn, J. E., *Ibid.*, **67**, 112 (1945).
[6] Audus, L. J., and Quastel, J. H., *Nature*, **159**, 320 (1947).
[7] Veldstra, H., and Havinga, E., *Rec. trav. chim.*, **62**, 841 (1943).
[8] Leopold, A. C., and Thimann, K. V., *Am. J. Bot.*, **36**, No. 4, 342 (1949).
[9] Thimann, K. V., and Schneider, C. L., *Ibid.*, **25**, No. 8, 627 (1938).

DEVELOPMENTAL ANALYSIS OF GENETIC DIFFERENCES IN PIGMENTATION IN THE AXOLOTL

BY H. CLARK DALTON

DEPARTMENT OF GENETICS, CARNEGIE INSTITUTION OF WASHINGTON, COLD SPRING HARBOR, N. Y.

Communicated by M. Demerec, April 4, 1949

The pattern differences in pigmentation exhibited by the white and black strains of the Mexican axolotl, *Siredon mexicanum*, were demonstrated by Haecker[1] in 1907 to be due to a single pair of alleles showing Mendelian segregation of the white condition as the recessive homozygote. Investigations on the nature of gene action in this case have been limited. Woronzowa,[2, 3] implanting pituitaries in postlarval stages, observed increases in pigmentation, and concluded from the differences in dosage required by the two strains that the recessive gene when homozygous raises the threshold of the skin to pituitary stimulation. Some information on the early developmental action of this gene has come from embryonic transplantations. The results of these experiments, summarized by DuShane,[4] have been interpreted as indicating that the genetic pigmentation differences between the two strains are mediated through the epidermis, which in the white genotype fails to provide some diffusible substance (oxidase or chromogen) necessary for melanin synthesis. The experiments here reported investigate further the developmental basis for this genetic difference in pigment pattern, by methods of tissue culture and embryonic transplantation.

Explantation of Neural Fold.—The pigment-forming potentialities of melanophores of the two genotypes, *DD* (black strain) and *dd* (white strain), were tested by growing explants of neural fold from the posterior trunk region in Holtfreter's solution in hanging-drop preparations (method of Twitty[5]). The cultures were maintained in a constant-temperature room at 18°C. The results have bearing on two questions: (1) Is the effect of gene *d* on the melanophores themselves or on their surroundings in the embryo? (2) Are the melanoblasts dependent on the epidermis for a substance necessary for melanin synthesis? Cultures of both genotypes showed the same characteristics of development. Melanoblasts began migrating out from the explants onto the cover glass in two or three days. In five days many cells had assumed the branched form characteristic of chromatophores, and some cells appeared full of gray granules. During the next four or five days more and more cells appeared full of granules, which darkened to become the black melanin granules found in differentiated melanophores. The development of melanophores in Holtfreter's solution is limited by the amount of nutrient reserves within

the cells. With the disappearance in about two weeks of the yolk platelets, which are large and numerous in early melanoblasts, development seemed to stop, although some cultures remained apparently healthy for as long as thirty-eight days. The developing time was long enough, however, to allow pigment production to the extent that many of the cells became very dark gray, and some black. Xanthophores were discernible, in cultures of both genotypes, one or two days after the melanophores appeared full of granules. No difference was found between cells of white and black strains with respect to extent of outgrowth or degree of pigmentation attained by melanophores. This result gives direct support to the idea, previously drawn from transplantation experiments, that the gene *d* acts on the environment of the chromatophores rather than on the chromatophores themselves. Since melanophores of both genotypes produced pigment in their inorganic medium, there is in this experiment no evidence that the melanoblasts of this species are dependent on a substance from the epidermis for melanin synthesis.

Embryonic Transplantations.—Information about the relationships between developing chromatophores and their surrounding tissues has been obtained by arranging these components of the two genotypes in different combinations. An experiment was planned to answer the question, do pigment cells of the same origin differ in development when in association with tissues of the black strain and of the white strain? Neural folds from the posterior trunk region of donors corresponding to Harrison stages 18–20 were transplanted to the midventral belly region of hosts in stages 22–23. Each donor provided one transplant in a black host and one in a white host, allowing subsequent comparison of graft chromatophores of identical genotype and age developing in the two environments. This transplantation site has the advantage of obviating any confusion of graft chromatophores with those of the host, since the ventral belly region does not become pigmented in either strain during the period of observation of this experiment. Six pairs of such transplants were made, three from black donors and three from white donors, with one extra case of a black-to-white graft. In all cases melanoblasts migrated out from the graft and became associated with host tissues. Two paths of migration were available to them; since the operation resulted in all cases in adhesions between skin and gut at the graft site, migrating cells could move out from the graft on either the skin side or the gut side. Pigment appeared in graft melanophores after 4 days in twelve cases and after 3 days in one case of black-to-white. Thus no delay was evident in the time of pigmentation of cells growing in the white hosts. With regard to degree of pigmentation, all melanophores became completely black in both types of host. The only apparent difference observed was that the melanophores of black origin in all cases appeared darker for a few

days in the white hosts than in the black hosts. This distinction was temporary, and does not necessarily reflect differences in the amount of pigment present, since variations in shape and orientation of the cells might easily give this appearance. From these direct comparisons it must be concluded that melanophores growing in association with tissues of the white strain reflect no disadvantage for melanin synthesis, in terms of differences in rate or degree of pigment formation, when compared with cells of the same origin growing in hosts of the black strain.

Comparison of the extent of migration of melanophores in the two types of host, on the other hand, revealed a striking difference. It was apparent from examination of the living animals that from all six grafts in black

TABLE 1

AREA IN SQUARE MILLIMETERS OCCUPIED BY MELANOPHORES FROM PAIRED BELLY GRAFTS

	BLACK HOST	WHITE HOST
Black donor	4.23	0.91
	1.35	0.59
	1.51	0.40
	..	0.37
White donor	3.70	0.77
	8.73	..
	4.01	0.52

TABLE 2

TOTAL NUMBER OF MELANOPHORES FROM PAIRED BELLY GRAFTS

	BLACK HOST	WHITE HOST
Black donor	452	119
	143	124
	57	76
	...	52
White donor	258	108
	341	...
	248	89

hosts melanophores had moved out extensively in both skin and gut tissues, whereas in four of the six grafts on white hosts the melanophores were restricted to the area of the graft itself and to the gut, none migrating into the skin. To obtain a more precise measure of the relative migration in the two types of host the following procedure was carried out. The animals were fixed in Zenker's fluid, four at 13 days and eight at 26 days after the operation. The skin was dissected off from the ventral side, cleared, and mounted in balsam; and the gut, from the stomach to the cloaca, was dissected out for observation in alcohol in a Syracuse dish. Camera lucida drawings were made of the areas occupied by donor melanophores in the skin and on the gut. Two drawings were made of each gut, to include both

sides. Donor melanophores migrating on the gut of black hosts surrounded the gut, in all six cases, and covered extensive areas anterior and posterior to the graft, whereas donor melanophores migrating on the gut of white hosts were in all six cases restricted to the side of the gut contacting the graft, and covered less area. The areas occupied by melanophores in the drawings were measured with a planimeter, and the figures converted to square millimeters. No account was taken of the curvature of the gut surface, the whole procedure being at best an approximation. It is apparent from Table 1 that the total area occupied by graft melanophores, regardless of dónor type, is much greater in black hosts than in white. The average area covered in black hosts was 3.92 mm.^{2}, compared with 0.59 mm.^{2} in white hosts. The smaller areas observed in white hosts might have been due to reduced numbers of melanophores as well as to impairment of their migratory activities. To check this possibility, melanophore counts were made in the skin and gut preparations. Results are presented in table 2, where it is evident that in all but one case more melanophores developed from grafts in black hosts than from corresponding grafts in white hosts. The difference in areas cannot be accounted for, however, by the relative numbers of cells from the grafts. When one calculates the number of pigment cells per mm.^{2} in each case, it is obvious that the melanophores in the white hosts are considerably more crowded than those in the black hosts. The average count for donor melanophores on black hosts was 70 cells per mm.^{2}, compared with 164 cells per mm.^{2} on the white hosts. The results are shown graphically in figure 1. These measurements lead to the conclusion that the melanoblasts in white hosts are subject to some inhibition of movement.

The migratory inhibition observed in the foregoing experiment was manifest in a region of the body which in the white strain never becomes pigmented and in the black strain is pigmented only after considerable delay in comparison with other areas. It may well be asked, therefore, whether there are special conditions in the ventral belly region which influence pigmentation there but have nothing to do with development of the normal pigment pattern dorsally. To test the development of melanophores of each genotype growing in a tissue environment of the other at the normal position in the body, reciprocal neural crest grafts were made between pairs of the two types of embryos in stages 23–27. Transplantation of the neural crest necessarily involves transfer also of the dorsal epidermis overlying it. In order to obtain graft melanophores growing in complete absence of any donor epidermis—a factor of some importance if there is a question of diffusible substances—the operation was made in two steps. First, for example, a piece of dorsal epidermis overlying the neural crest of a black embryo was replaced by a piece of flank epidermis from a white embryo. Flank, rather than dorsal, epidermis was chosen to

prevent the possibility of transplanting any adhering neural crest cells. On the following day the black neural crest, covered now by white epidermis, was transplanted to a white embryo. Thus, donor melanophores could differentiate in association with epidermis of host type only. By this method five transplants of black neural crest to white hosts and four transplants of white neural crest to black hosts were made. For comparison, a few neural crest transplants bearing donor-type epidermis were also made (two black to white, three white to black). The experiment is not yet terminated, since eight of the animals are still growing, but the effect on larval pigmentation was clear and consistent for each type of transplant.

When black neural crest covered by white flank epidermis was grafted

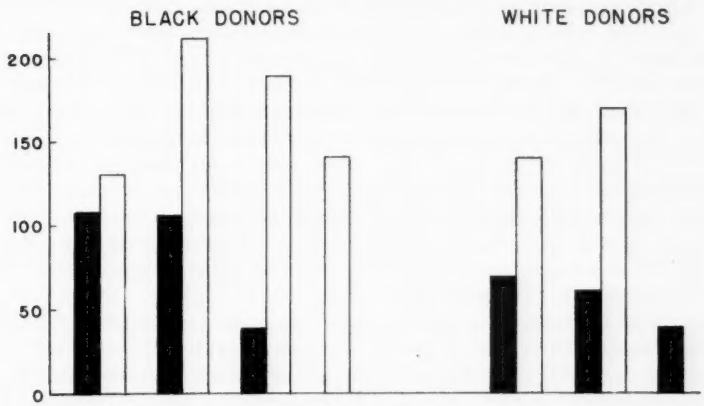


FIGURE 1

Number of melanophores per square millimeter from paired belly grafts in black hosts (solid) and in white hosts (unshaded). Adjacent columns represent grafts from the same donor.

to a white host, a dense black line of melanophores appeared beneath the epidermis at the edge of the graft region, which was clearly indicated by a discrepancy in the dorsal fin in that area. When black epidermis was grafted along with the neural crest to a white host, the dorsal fin developed intact and graft melanophores migrated abundantly into the donor portion of the fin but not into the adjacent host portion of the fin. The ventral limit of donor melanophores was not a sharp border, as in the cases described above, but the few large melanophores appearing below the graft remained on the dorsal part of the flank. These results suggest that the mobility of *DD* melanoblasts is greater in an environment of *DD* tissue than in one of *dd* tissue.

When white neural crest covered by black flank epidermis was grafted to a black host, the graft region was later indicated, as in the reciprocal experiment, by a discrepancy in the dorsal fin. Donor melanophores appeared widely dispersed on the flank under the graft, and in two cases in extensive areas posterior to the graft region as well. These melanophores appeared somewhat smaller than the host melanophores, and did not form by overlapping the large irregular spots characteristic of the host pigment pattern at this stage. A further indication that the area of the flank beneath the grafts was inhabited by donor chromatophores came from the xanthophores, which in the white animals used in these experiments were much deeper yellow than those of the black animals. This distinction was very striking in black hosts bearing white neural crest grafts, the deep yellow areas occupied by donor xanthophores being visible to the naked eye. The donor melanophores and xanthophores migrating from grafts of white neural crest covered by donor epidermis on black animals showed the same wide dispersal on the flank of the host as was noted in the transplantations just described. These results suggest that the mobility of *dd* melanoblasts also is greater in *DD* tissues than it is in *dd* tissue. The evidence from all these homotopic neural crest transplantations indicates that the phenomenon observed in the belly grafts is not a special case, but that conditions affecting melanophore migration in the normal environment of the developing pigment cells are different in the two genotypes and are significant in the formation of the genetic pigment patterns characteristic of the two strains.

Discussion.—On the basis of the hypothesis that the gene *d* affects the epidermis of the recessive homozygote in such a way that it does not supply to the underlying melanoblasts a substance necessary for pigment production, which is, however, furnished by epidermis containing the allele *D*, one would have expected that *in vitro* in Holtfreter's solution melanoblasts of both genotypes would fail to produce melanin. That they both did become pigmented argues against this hypothesis. Some epidermal cells were undoubtedly explanted with the neural folds, but if these were responsible for allowing pigment formation in the melanophores, there should have been a difference in this respect between explants of the two strains, whereas no such difference was observed. To account for the presence of pigment cells in normal white animals, DuShane[6] suggested that there may be some melanophores that are independent of the epidermis effect. If the melanophores that developed from explants of white neural fold represented such independent pigment cells, one would have expected to find in addition many melanophores that remained colorless. A few non-pigmented cells otherwise morphologically resembling chromatophores were observed, but they were few in number (less than 4%) and occurred with the same frequency in explants

from both strains. The idea of diffusible substances from the epidermis as the mechanism of gene action in this case is also inapplicable to the results of the transplantation experiments, where melanin synthesis was carried out equally well by melanoblasts in both types of tissue environment. This hypothesis concerning the nature of the epidermis effect is based on the assumption that melanoblasts migrate in the white axolotl just as they do in the black strain and in other pigmented species that have been used for reciprocal transplantations. This assumption does not explain the larval pattern distribution of melanophores in the white strain; furthermore, the impairment of mobility of melanoblasts in *dd* tissues reported here suggests that it may not be valid. The results of the present transplantation experiments, as well as the larval pigment pattern differences between the white and black strains of axolotl, are easily explained by the hypothesis that the melanoblasts of both genotypes are capable of producing pigment but that a reaction between melanoblasts and surrounding tissues of the white strain inhibits melanoblast migration. Whether this reaction is physical or chemical in nature cannot be stated at this time. Further experiments to obtain critical evidence on the problem are in progress.

Summary.—The differentiation of pigment cells from explanted neural folds of black and white axolotls in Holtfreter's solution has shown that potential melanophores of both genotypes are capable of melanin synthesis *in vitro*. Melanophores from grafts of neural folds to the midventral belly region of both dark and white hosts exhibit no disadvantages for melanin synthesis in white embryos. Chromatophores from reciprocal grafts of neural crest between white and dark strains also produce pigment in both types of host. A striking difference in the distribution of melanophores according to host genotype has appeared in both homotopic and heterotopic grafts. Melanoblasts of both genotypes appear to migrate freely beneath epidermis of the black strain but not beneath epidermis of the white strain. The results of all experiments support the idea that the genetic differences in pigmentation are mediated through differences in the tissue environment and not in the chromatophores themselves, the white pattern resulting from an inhibitory effect concerned not with conditions for melanin synthesis but with the migration of melanoblasts.

I wish to express my appreciation to Dr. E. W. Caspari for his continuous advice and encouragement during the course of this investigation.

[1] Haecker, V., *Zool. Anz.*, **31,** 99–102 (1907).
[2] Woronzowa, M. A., *Trans. Lab. Exp. Biol. Zoopark Moscow*, **4,** 107–124 (1928).
[3] Woronzowa, M. A., *Arch. Entw. Mech. Org.*, **115,** 93–109 (1929).
[4] DuShane, G. P., *Spec. Pub. N. Y. Acad. Sci.*, **4,** 1–14 (1948).
[5] Twitty, V. C., *J. Exp. Zool.*, **100,** 141–178 (1945).
[6] DuShane, G. P., *Ibid.*, **72,** 1–31 (1935).

INHERITANCE OF CAROTENOID DIFFERENCES IN LYCOPERSICON ESCULENTUM STRAINS

BY G. MACKINNEY AND J. A. JENKINS

DIVISIONS OF FOOD TECHNOLOGY AND GENETICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA

Communicated by E. B. Babcock, April 1, 1949

Two recessive mutant types of the common red-fruited tomato have long been known, the tangerine and the yellow. MacArthur[1] established the fact that these two color variants are due to recessive alleles at two loci situated on different chromosomes. They are therefore inherited independently. Using MacArthur's symbols, the genotypic formulae of the three different color types with respect to the two loci are: red *RRTT*, yellow *rrTT* and tangerine *RRtt*.

Ultimately, we hope to evaluate quantitatively the various combinations, such as, for example, those producing red fruit, namely *RRTT*, *RRTt*, *RrTT* and *RrTt*. The precise rôle of *R* and *T* genes cannot be established until the double recessive *rrtt* has been described. So far as we are aware, this has not yet been described, though we hope to identify it shortly.

In this paper we present evidence identifying the major carotenoid components for the three homozygous lines, and the relative proportions in which they are present. Data assembled over the past three seasons confirm and extend much that has been published[2-4] though some discrepancies exist. Pigmentation in a given phenotype is the expression of two factors, genetic constitution and environment. Lycopene formation is particularly sensitive to temperature. Thus, the fact that some twenty pure lines of yellow tomatoes studied here (differing greatly in pigment content and in other respects) invariably contained small quantities of lycopene, does not invalidate the argument[2] that the *R* gene is required (*inter alia*) for its normal production. In the absence of *R* the major and normal line of synthesis is blocked, but this need not preclude the feasibility of lycopene formation, under certain environmental conditions, in small amount, possibly with greater difficulty, by another path.

The red tomato owes its characteristic color to lycopene, the orange-colored tangerine to prolycopene, as noted by Zechmeister *et al.*,[3] and the yellow tomato to β-carotene and similar components. The three types are so characteristic that an examination of the absorption spectrum of a crude extract suffices to determine the type. Characteristic absorption curves are shown in figure 1. They are remarkably constant in shape for a given strain.

It is particularly interesting that while prolycopene determines the color of the tangerine tomato, its absorption is dominated (curves IV and V, figure 1) by pigments with maxima between 430 and 400 mμ. This is

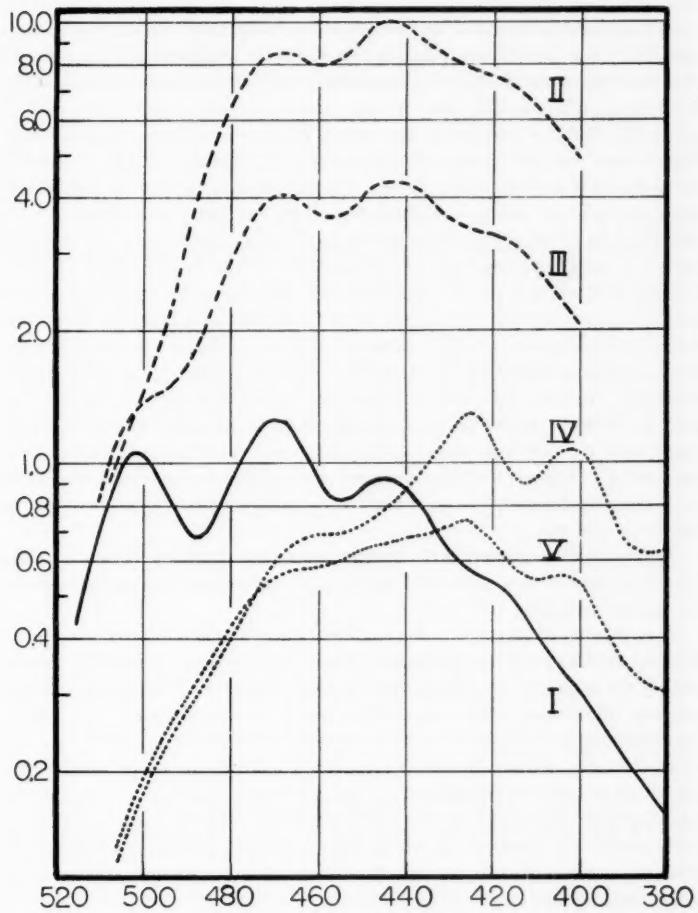


FIGURE 1

Absorption curves for typical extracts. Ordinate: optical density in arbitrary units; abscissa: wave-length. Curve 1 red, *RRTT;* curves II and III yellow, *rrTT*, curve III has more lycopene than curve II; curves IV and V tangerine *RRtt*.

accounted for by two pigments, all-trans ζ-carotene, and the "unidentified carotene I" of Porter and Zscheile.[4] For convenience in reference, it is tentatively designated as a poly-cis ψ-carotene. It has an all-trans spectrum apparently identical with Porter and Zscheile's "unidentified carotene II,"[4] but there is some reason for doubting whether their carotene II is a member of the ψ-set. Traces of other cis ψ-carotenes are present in the tangerine, and possibly also in the yellow tomato. They are readily separable, but are present in the yellow tomato in minute quantity, sufficient only for comparison of iodine-catalyzed spectra, indistinguishable from that for poly-cis ψ-carotene. The absorption curves for poly-cis ψ-carotene and its iodine equilibrium mixture are shown in figure 2. As surmized by Porter and Zscheile,[4] it cannot be a poly-cis lycopene. In so far as pure lines homozygous with respect to the *R* and *T* loci are concerned, poly-cis ψ-carotene appears to be unique to the tangerine. It is found in quantity (cf. [4, 5]) on the Tswett column as a bright yellow zone below prolycopene, above ζ-carotene. It is therefore spectroscopically and chromatographically comparable with "carotene I" of Porter and Zscheile. But for different spectroscopic results reported by Zechmeister and Le Rosen,[5] we should also consider it to be identical with zone 16 of their table 2. Attempts to crystallize poly-cis ψ-carotene have not yet been successful. Repeated chromatograms have failed to eliminate the shelf, *ca* 455 mμ in the poly-cis spectrum, or to change its height relative to the height at 430 mμ.

*Experimental.—Materials:** Twenty-two pure lines of yellow tomato, six of tangerine and three red have been examined, several repeatedly, to assess variability.

Preparation of Extracts: To minimize possible change, fruits were extracted rapidly, and measurements made immediately. Normally about 100 g. fresh weight of tomato from a single plant was blended in 300 ml. acetone, filtered on a Buchner and the residue re-suspended and the blending repeated.

In the case of the yellow and tangerine tomatoes the fleshy residue was white after a second filtration, but the skins were yellow if non-carotenoid alkali soluble skin-pigment was present. Where the red tomato contained carotenoid skin pigment, grinding with sand and a small quantity of petroleum ether was also necessary. Three such treatments gave a fairly colorless residue.

Successive portions of the aqueous acetone extract were then transferred to 100 ml. petroleum ether, and water added to remove acetone, until all pigment had been transferred to the petroleum ether.

The petroleum ether solution was then freed from acetone by washing, and dried with anhydrous sodium sulfate.

Absorption Spectra: The absorption spectra of suitably diluted aliquots

in petroleum ether were immediately determined on the Beckman DU spectrophotometer, tungsten lamp, in the range 520 to 330 mμ at 5-mμ intervals.

The curves shown in figure 1 are typical of more than 100 analyses. Apart from the absolute quantity of pigment produced, a given yellow strain shows greatest variability with respect to its lycopene content. This

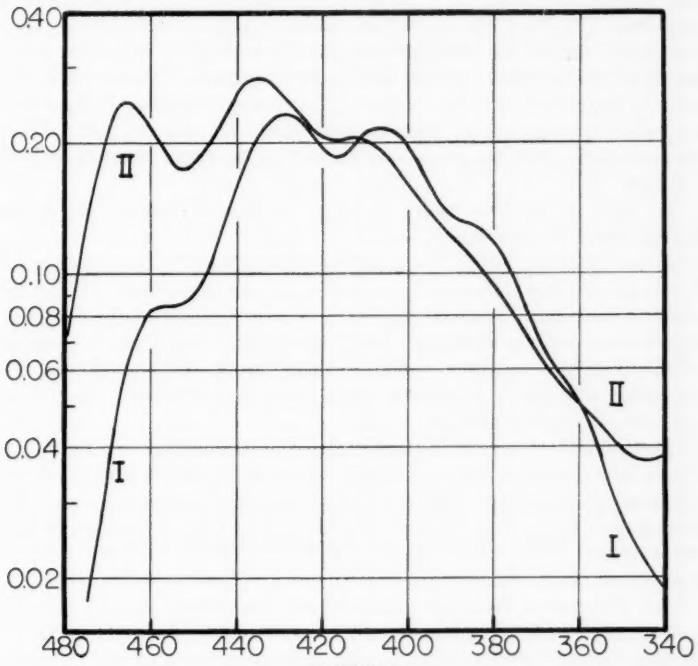


FIGURE 2

Absorption curves for ψ-carotene before (curve I) and after (curve II) iodine catalysis. The curves are directly comparable with respect to concentration, and therefore to optical densities.

is reflected in the absorption at 500 mμ, relative to other wave-lengths, illustrated in curves II and III of figure 1. They were obtained from extracts of tomatoes taken from the same vine, at the same time. Those represented in curve III had pale pinkish streaks. Lycopene was also detected in the extract represented in curve II, but in lesser amount.

Red tomatoes varied in carotene content but the spectrum was invariably dominated by lycopene (curve I). Tangerine extracts fell into two cate-

gories, consistently reproducible over the period during which fruit could be harvested. Curves IV and V represent extremes in the strains of this line.

The Individual Components: Petroleum ether extracts, freshly prepared as above, were chromatographed on magnesium oxide-hyflo-super cel (1:2). Columns packed to a height of 25 cm. and 2 cm. diameter were used. The more weakly adsorbed members were separated with this solvent, and identified. For strongly adsorbed components, fresh columns of adsorbent were used, the solvent being petroleum ether containing from 5 up to as much as 15% acetone. Before subsequent spectroscopic examination, the acetone was removed, in the usual way. As recommended by Zechmeister, an iodine catalysis was routinely included in the run, the two sets of spectrophotometric data being made simultaneously on aliquots from the parent solution.

In view of available detailed reports[2-5], brief comment will suffice except where discrepancies exist.

Yellow Tomatoes: Phytofluene, barely detectable in trace amounts, small bands at 370 and 348 being just discerned; the β-carotene set showed maxima unchanged by iodine catalysis; ζ-carotene could not be detected; minor components with well-separated bands follow next, giving a common iodine-catalyzed spectrum indistinguishable from that for poly-cis ψ-carotene; all-trans lycopene was noted, max. 505, 470, 445; shifting to 500, 468, 445 mμ on addition of iodine.

Tangerine Tomatoes: The detail reported by Zechmeister and Le Rosen[5] (table 2) is confirmed except for zones 15 and 16, neither of which give the lycopene spectrum with iodine. Our results are shown in figure 2. We doubt also whether zone 7 of their report is a true lycopene isomer, though we have not been able to free this clearly defined yellow zone from a component which, with iodine, gives a small maximum at 498 mμ.

Red Tomatoes: These typically contain phytofluene and all-trans β-carotene. ζ-carotene has not been unequivocally detected in *RRTT* strains. Traces of γ-carotene, and predominantly all-trans lycopene complete the pigment complex.

In all cases, above the lycopene fractions, a yellow-orange zone is discerned. Sometimes this shows xanthophyll characteristics and spectrum. More often the spectrum is ill-defined, and recovery by elution is poor. In no case does this comprise a significant fraction.

Relative Abundance of Components: Extracts were prepared, as described above, from composite samples involving several plants of a given line, and by using larger quantities of fruit, up to 5 lb., extracts could be transferred to petroleum ether and made relatively concentrated in pigment without heating in any manner, so that for spectroscopic work, dilutions of 1 to 100 or 1 to 200 were necessary. Such concentrated extracts gave

best results for this phase. Aliquots (usually 5 ml.) of the concentrated solution were adsorbed on a column. The absorption spectrum of the properly diluted original solution was immediately determined and an optical density calculated for the concentrated solution (column 2, table 1, etc.). The absorption for the individual components was then determined, in successive columns, and the similarity between the sum of the individual fractions (column 3) and the original (column 2) indicates the extent to which the absorption has been accounted for. This is of particular interest in the tangerine tomato where the absorption indicates the prevalence of pigments of low tinctorial power with high absorption below the main absorption for prolycopene.

Because speed in separation was essential, as many as three adsorbent columns were used, the chromatograms being developed from the start

TABLE 1

YELLOW TOMATO: OPTICAL DENSITIES OF EXTRACT AND OF FRACTIONS

WAVE-LENGTH	ORIGINAL	SUM OF FRACTIONS	PHYTO-FLUENE	β-CAROTENE SET	MISCEL-LANEOUS*	LYCOPENE* (IMPURE)
500	1.60	1.21	..	0.6	0.15	0.46
490	3.12	2.46	..	1.56	0.45	0.45
480	6.10	4.42	..	2.53	1.22	0.67
465	8.38	5.08	..	2.50	1.70	0.88
445	9.41	6.75	..	3.07	1.95	1.16
425	7.87	4.93	..	2.30	1.50	1.13
400	5.00	3.05	..	1.31	0.92	0.82
370	1.80	1.11	0.20	0.25	0.34	0.32
350	1.19	0.86	0.21	0.24	0.17	0.24

* The miscellaneous group includes minor components and a larger yellow fraction immediately below the lycopene. Both yellow and lycopene fractions require several chromatograms for spectroscopic purity. A well-defined maximum for the lycopene exists *ca* 505 mμ.

with solvent containing different percentages of acetone (usually 0–2, 5 and 10, though solvent with as much as 15% acetone has been used).

There is a two-fold advantage in reporting the data in terms of optical densities at this stage. It avoids arbitrary compromises in the value selected for the absorption constants. Thus components in table 2 designated L_1, L_2 and L_3 are undoubtedly lycopene-*like* isomers. They have been crystallized, but the curves differ, even after iodine catalysis, and also from the curves obtained prior to crystallization (the change being due to partial isomerization during concentration).

Secondly it enables a comparison to be made between the original optical density, column 2, and the sum of the components, column 3. Losses for more strongly adsorbed components account in part for the differences. This procedure is useful in showing whether there may be a significant

contribution by an undetected pigment, and whether unsuspected isomerization has occurred to any extent during evaluation.

The wave-lengths for which optical densities are reported in the tables were chosen on the basis of the characteristics of the absorption curves for the originals, cf. figure 1. Routine measurements were made throughout the whole range, however, without detection of abnormalities.

TABLE 2

TANGERINE TOMATO: OPTICAL DENSITIES OF EXTRACT AND OF FRACTIONS

WAVE-LENGTH	ORIGINAL	SUM OF FRACTIONS	PHYTOFLUENE	β-SET	ζ-CAROTENE	CIS ψ-CAROTENE	PROLYCOPENE	YELLOW ZONE	L_3*	L_2	L_1	X
490	30.1	25.4	..	3.0	..	0.7	11.2	1.0	1.9	2.8	4.5	0.8
480	46.5	41.2	..	5.9	..	1.2	18.3	1.7	3.5	3.5	5.1	2
465	84.8	64.1	..	5.9	<1.0	4.8	27.5	3.9	6.2	5.0	7.0	2.8
435	127.5	121.3	..	5.8	20.6	14.2	56.5	4.2	7.0	4.8	6.0	2.2
425	156.3	144.4	..	4.3	53.1	15.6	50.0	3.3	5.5	4.1	5.2	2.3
410	117.5	101.2	..	2.9	27.1	14.1	40.4	3.1	5.0	3.3	4.3	2
400	125	119.5	0.2	2.2	53.3	13.8	35.0	2.4	4.8	2.7	4.8	1.8
370	68.1	61.4	12.1	1.3	21.3	5.4	15	...	2.0	1.8	2.5	...
360	51.3	47.2	8.8	0.5	16.0	4.0	11.6	...	2.0	1.8	2.5	...
350	43.8	43.5	15.6	0.4	11.9	2.3	8.6	...	1.6	1.4	1.7	...

* The designations L_1, L_2 and L_3 represent lycopene isomers separated by a small yellow zone from prolycopene. X is a xanthophyll fraction. L_3 contains several minor zones.

TABLE 3

RED TOMATO: OPTICAL DENSITIES OF EXTRACT AND OF FRACTIONS*

WAVE-LENGTH	ORIGINAL	SUM OF FRACTIONS	PHYTOFLUENE	β-CAROTENE	LYCOPENE
500	117	103.1	...	2.3	100.8
495	97.9	84.2	...	3.2	81.0
480	101.2	90.0	...	5.2	84.8
470	140.6	119.6	...	5.0	114.6
445	102.8	85.3	...	5.4	79.9
425	63.7	50.9	...	4.3	46.6
400	33.7	25.7	0.2	2.3	23.2
370	18.9	15.7	5.2	0.4	10.1
360	17.1	14.8	3.2	0.3	11.3
350	16.9	15.3	6.1	0.4	8.8

* γ-Carotene and xanthophyll are minor contributors, *ca* 0.5 at their maxima. The lycopene value is low owing to incomplete elution.

With the reservation that because of the necessity of rapid separation, the zones are not highly purified from a spectroscopic standpoint, the values reported in tables 1, 2 and 3 are a sound measure of the relative abundance of the various components.

The data for tables 1, 2 and 3 are based on equivalent weights of fresh fruit, so that a rough comparison of absolute values is possible. The various yellow lines differ greatly in the amount of pigment produced,

and there may be some significant variations with respect to individual components. This phase must be deferred at present. Based on a comparison of table 1 with tables 2 and 3, it appears with respect to all lines of composition *rrTT* that they are markedly deficient in their ability to produce carotenoid pigment; that no component has yet been identified as unique to the yellow tomato; and finally that no component has been found in greater absolute amount in the yellow than in tangerine or in red lines.

When table 2 and table 3 are compared, the total pigment production for tangerine and red lines is approximately equivalent.

As shown in table 1, β-carotene is clearly a major component in the yellow lines and lycopene is present in small amounts. In the red (table 3), the absolute amount of β-carotene is somewhat greater but the proportion, relative to other components, is greatly reduced. The overwhelming contribution is made by all-trans lycopene. The tangerine (table 2) is comparable with the red, so far as β-carotene is concerned, higher in phytofluene, and particularly high in all-trans ζ-carotene, poly-cis ψ-carotene and prolycopene. It will clearly be of interest to isolate the heterozygous genotype *Rrtt* for comparison, with respect to these components.

The evidence indicates certain tentative conclusions as follows:

In the absence of *R*, the *T* gene is responsible for pigment production on a limited scale, particularly with respect to lycopene, cf. reference 2, p. 242. In the absence of *T*, the *R* gene is responsible for large quantities of ζ-, poly-cis ψ-carotenes and prolycopene. In the presence of *R* and *T*, these pigments or their immediate precursors are converted to lycopene. The suggestion of Porter and Zscheile (reference 4, p. 544) that ζ-carotene may be concerned in lycopene formation seems therefore entirely plausible. The open-chain structure recently postulated for ζ-carotene[6] also supports this view.

* Lines studied in detail include the red "Bonny Best," a tangerine from Dr. J. W. Lesley and the yellows "Snowball" from Good Luck Gardens and MacArthur's tester stock 902.

[1] MacArthur, J. W., *J. of Genetics*, **29**, 123 (1934).

[2] Le Rosen, A. L., Went, F. W., and Zechmeister, L., *Proc. Nat. Acad. Sci.*, **27**, 236 (1941).

[3] Zechmeister, L., Le Rosen, A. L., Went, F. W., and Pauling, L., *Ibid.*, **27**, 468 (1941).

[4] Porter, J. W., and Zscheile, F. P., *Arch. Biochem.*, **10**, 537 (1947).

[5] Zechmeister, L., and Le Rosen, A. L., *J.A.C.S.*, **64**, 1075 (1942).

[6] Nash, H. A., Quackenbush, F. W., and Porter, J. W., *Ibid.*, **70**, 3613 (1948).

INFRARED ABSORPTION IN FIELD STUDIES OF OLFACTION IN HONEYBEES*

BY WALTER R. MILES AND LLOYD H. BECK

YALE UNIVERSITY SCHOOL OF MEDICINE AND DEPARTMENT OF PSYCHOLOGY, YALE UNIVERSITY

Read before the Academy, November 19, 1947

Introduction.—Some two decades ago it was pointed out by Sherrington that the biological function of a specialized sense organ, such as the eye or ear, is to lower the threshold of the nerve fiber to some one form of energy. The senses may be classified in terms of the energy form which serves as the adequate stimulus for each modality. Each sense mechanism that plays a rôle in the integrated behavior of an organism appears subject to stimulation by mechanical energy, chemical energy or radiant energy. According to this widely accepted scheme of classification hearing, touch, pressure, equilibrium sense and the lateral-line organ sense of fishes are all mechano-receptors; vision, the sense of heat or warmth, and the sense of cold are radio-receptors; while the chemo-receptors are smell, taste, common chemical sense and the vomero-nasal organ sense. The theory of olfactory stimulation proposed by the authors[1] would, if substantiated, suggest that the sense of smell be reclassified and placed among the radio-receptors.

In order to test the validity of the hypothesis that stimulation of the olfactory receptors results when the radiation of heat from smell receptors is suddenly increased, a wide range of new investigation is required. Some of the more simply arranged of these desired experimental observations have been undertaken by us, and the present report presents a part of this work, specifically some accomplished with the honeybee, *Apis mellifera* Linn.

Insects are peculiarly well adapted for behavioral studies on the sense of smell due to the peripheral placement of their olfactory receptors along their antennae rather than within a nasal cavity.[2, 3] The worker honeybee offers the additional advantages of high activity level, marked socialization and well-integrated exploratory and work habits. There is a rich scientific literature concerning the olfactory sense in bees. It has been shown that their sense of smell exhibits a marked parallelism with that of man. Von Frisch, as is well known, demonstrated that pairs of different chemical substances judged alike in odor by man were responded to very similarly by the bee. While it is interesting to attempt reinterpretation of older work from the standpoint of a radiation hypothesis, it is obvious that newly designed equipment and methods must be developed and new experimentation conducted in order to obtain critical results for appraisal.

Apparatus and Methods.—For our studies on insects we appeared to need

what may be called osmic radiation receivers. These may be thought of as artificial noses since they provide chambers which may be filled with odoriferous gases. The apparatus described by us in another paper[1] was designed for some work on cockroaches and is one type of osmic radiation receiver. It is a laboratory type adapted to a variety of insects. A second type of receiver based on the same general principles but considerably simplified has been developed for field studies with bees and other insects. Its general form is illustrated in the accompanying figures 1 and 2. The receiver is made of ordinary black iron pipe fittings as indicated in figure 3. The central part, *B*, is a black iron cross-tee with horizontal openings of one inch, top and bottom openings, three-quarter inch. The inside of this tee serves as a gas chamber and is sealed leak-tight. A black iron diaphragm at *C* closes off one end; *A* is a closed cavity for liquids at the bottom of the gas chamber (we may refer to this as the "perfume bottle"); *F* is a cavity, an extension of the air-tight system of the gas chamber, which is filled with calcium chloride, anhydrous ($CaCl_2$) to absorb water vapor. The absorbent is separated from the gas chamber by a number eight mesh wire screen. *D*, represents the infra-red passing window (described more fully below) which sealed off the front of chamber *B*. Finally *E* in figure 3 represents an accessory chamber for temperature control. The bottom side of this chamber is bounded by the black iron membrane brazed to the iron cross-tee at position *C* in the figure. Ordinarily a small amount of dry ice is placed in receptacle *E* to lower the temperature of the gas chamber below that of the environment. The entire iron unit stands eight inches high, six inches deep, from *D*, as its front, and three inches wide; its weight is about seven pounds. By means of a railing flange it is fastened to a wooden base with two screws. The box was made from old well-weathered chestnut, as exhibited in figure 2. A pair of such receivers was used in our field experiments. One unit was "active" in the sense that infra-red could pass into it. The other unit was used as control. Directly behind the thallium bromide iodide window in the control unit was a glass disc. Thus infra-red rays could not penetrate into the gas chamber of the control unit but were reflected back by the glass.

The window in the front of each osmic radiation receiver was made of a crystal of thallium bromide iodide designated KRS-5.[4] Sections were sawed from the crystal and then polished to a thickness of 2.0 mm. These discs were 50 mm. in diameter.[5] By means of "Quartz Cement"[6] they were sealed into recessed grooves that had been cut in the iron collars at the front of each unit. The groove for the control unit was cut deep enough to receive first a glass disc and then the KRS-5 disc next the glass and in front of it. The thallium bromide iodide was therefore the biological contact surface from the outside for both units.[7]

A flame should not be used against KRS-5 as the fumes from it are very

toxic. When polishing this material rubber gloves and a mask should be used.[8] Sections of this crystal take and retain a good polish. This is one of the great advantages it possesses over calcium chloride or potassium bromide crystals both of which require frequent repolishing. The polished KRS-5 surface is not observably toxic to bare fingers or to insects in contact with it.

The second and more important advantage possessed by the KRS-5 crystal is its transmission characteristics for infra-red radiation as revealed in figure 4. Beginning at 0.6 μ (micron) 600 mμ (millimicrons) or 6000 A (Angstrom units), which is in the region of the yellow band of the visible spectrum, the transmission rises precepitately to 40 per cent. It increases to about 55 per cent for the extreme visible red, 0.7 μ, and then somewhat more slowly rises in the near infra-red reaching 65 per cent for 1 μ (10,000 A) and at 1.5 μ levels off at 70 per cent which is maintained evenly to an extreme of 31–32 μ. The heat radiation in the case of human beings would probably cover the range 7.5 to 15 μ with a maximum at about 9 to 10 μ. The radiation from bees, judging from the similarity of their odor discriminations to those of human beings, probably covers at least a part of this 7.5–15 μ range. It is obvious that the KRS-5 provides an amply wide and high transmission range for all such experimental purposes.

A pair of osmic radiation receivers such as described above lends itself to various methods of use in the field. Both windows should have the same temperature which is lower than the environs. (1) The two windows may be exposed side by side for attracting bees directly. (2) The windows may be equipped with short tubes or platforms extending in front of them, bait may be placed here and bees artificially transferred to the two baited areas. This latter procedure has proved the more useful experimental plan and lends itself to the following four variations.

(*a*) Bait removal method: Bees transferred to baited tubes. Before the bait has been completely used up they are gently shaken from the tubes below the units and the bait tubes are removed. Predictions: Bees should return to the window areas. The window that now presents the bait odor (or more of it) should for a time collect the more bees on or near it.

(*b*) Bait exhaustion method: Bees transferred to baited tubes, the tubes are left undisturbed by the experimenter for 20 minutes or more. Predictions: After the bait in the tubes has been exhausted the window that continues to provide more bait odor should hold the attention of more bees, more bees should stay for a time in this tube or bees should more frequently return to this tube.

(*c*) Bait interchange method: Same as (*b*) above except that within one or two minutes after the bait has been exhausted the tubes should be exchanged in position between the units with the least possible disturbance of the few bees in them. Predictions: Bees in the tube placed in front of the experiment window should tend to remain and the bees that have been placed in front of control window should tend to return to the experiment window. If some bees have made trips from and to the bait a temporary position habit may interfere with the predicted result.

(*d*) Bait repetition method: Same as (*b*) above except that when the original bait is nearly exhausted a little more is added equally in each tube and this process is repeated from time to time to keep the bees working for one or more hours, making trips from the bait to the hive and back. Predictions: If the amount of available bait is kept minimal the window that supplies more bait odor should attract the larger population of bees. In work with (*d*) the right-left positions of the units should be exchanged at intervals of 20 or 30 minutes. The units may also be placed one on top of the other and this position changed in a chance order with the right-left positions. Other changes of orientation should be used such as rotation of the units by 180° or 90°.

If we set up our receiver units on a platform near an active hive of bees on some seasonable day when there is not too much natural nectar available, and if we wait long enough some diligent worker bee will find them. She may alight on the window which passes infrared and move slowly, waving her antennae as if near nectar, but she can't get it. She can't carry anything away but an experience. There is no tangible reward for the effort expended. She is unlikely to return to the hive and engage in a dance to attract other bees out to visit this honey smell. Some method of transient baiting to capture the bees' attention was required if experimentation was to proceed without unduly long delays. The method adopted involved two principal features, see figure 5. First, a short plastic tube was provided as an extension from the opening which surrounded each window. Usually three drops of honey were placed in each tube on filter paper or a bit of cloth attached inside and at the bottom of each tube. Second, to a wand, four feet long, filter paper was bound at one end and smeared with honey. This end of the wand was placed near the opening of a hive; after a few minutes the filter paper was well covered with bees. The wand was then lifted, brought back to the units and 10 or 12 bees, an equal number for each unit, were transferred to the bait tubes by gently rolling the wand in such a way as to walk the bees off into the tubes. Each experiment began thus by causing bees to work in the immediate vicinity of the two windows.

Results.—Experiment I: The unit containing the "active" or experimental window was on the left side, see figure 6; the control unit with the window that would not pass infra-red was on the right. The thermometer showed 70°F. for outside temperature. The date was September 28, 1947. The amount of natural nectar available was not great. The method employed was a combination of (*a*) and (*d*) above. After the bees had been transferred to the tubes, they cleaned up the available honey rapidly. Drops of honey were added, equally to each tube, until it was found that the bees were working continuously and well. Photo A was taken at 10:30 A. M. At 10:45 (Photo B) the honeybait seemed exhausted in both tubes, only one bee remained in *CW*, those in *EW* were rushing about. At 10:47 both tubes were removed, and placed in an air-tight container. When the tubes were removed, the bees were gently shaken out of them down near the ground below the units. One minute later (Photo C) ten

bees were found on *EW*, none on *CW* but three were flying about *CW*. This difference could occur by chance once in about 1000 times. At 10:49½ (Photo D) nine bees were on the "active" window and three flying about it. None were at the control window. At 10:54 (Photo E) three bees were on *EW*, and none on or near *CW*. At 11:09 (Photo F) three were around *EW*, no bees were near *CW*. During the next 15 minutes both windows were continuously observed, but no photographs were taken. Within this period 12 bees alighted on "active" window and eight paused in flight in front of it. No bees alighted on the control window, and only two paused in flight in front of it. This difference represents a chance of one in about 4000.

Experiment II: This illustrative experiment follows method (*c*). The tubes were left in place after the bait had diminished practically to the vanishing point judged by the behavior of the bees. Later when most of the bees were found in tube *EW*, the tubes were gently exchanged with the bees in them. This experiment was made in mid-October, 1947. At 9:35 A.M. measured honey bait was placed in the two tubes; no bees came on their own initiative. During the next few minutes bees were transferred equally. After 9:50 A.M. no more bees were transferred. The "active" window (*EW*) was now on the right side; the control unit (*CW*) on the left. At first the bees were working in equal numbers in the two tubes. By 9:57 most of them were in tube *EW;* only one was in front of *CW*. At 9:59 all were in front of *EW*, results to be expected by chance about once in 250 times. At 10:03 most all bees were in tube *EW*. At 10:07 all present were in *EW*. At 10:08 the tubes were gently exchanged and it was not observed that any of the bees left during the exchange. This maneuver brought the great majority of the bees in front of the control window. At 10:20 all were observed to have collected at the "active" (*EW*) window. At 10:21 the tubes were again gently exchanged. At 10:25 some had returned to the "active" window. At 10:34 most of them had returned. By 10:37 all bees present were in the tube in front of the "active" window. A total of 47 minutes had now elapsed since any bees had been transferred to the tubes; 30 minutes since the bait, as a tangible reward, had worn out. These results are typical for other experiments we have conducted.

Field experiments such as the two described above do not, as a rule, yield such definite results when there is abundant nectar to be gathered from wild or cultivated plants available to the bees. This general finding would appear predictable, since the reward the bees secure from their efforts in collaboration with our experiments is meager in comparison with what an equal amount of time devoted to abundant field bloom would give.

Results for experiments attempted in summer when the field supply of nectar was very plentiful lend themselves to presentation in tabular form. These tests were made on a farm in Otsego County, N. Y. in mid-August,

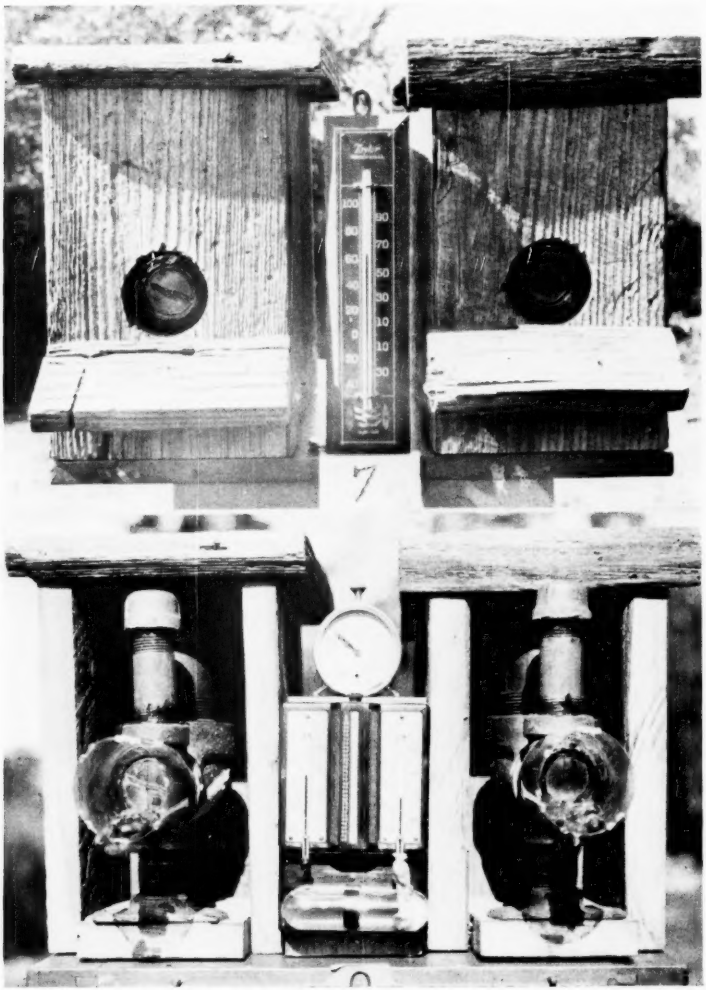
1948.[9] The bees were very busy on buckwheat and red clover, fields of both were lying within a quarter mile of the colonies studied. Method (*b*) was employed, i.e., after bees had been transferred to the baited tubes the situation remained undisturbed for 20 minutes or more. Tests made on the same day succeeded one another at short intervals. Our grouped data for eight tests are presented (see table 1) in terms of the numbers of bees counted at *CW* and at *EW*, 5, 10, 15, and 20 minutes, respectively,

FIGURE 1

A pair of osmic radiation receivers in their housings, placed side by side as used in some olfactory experiments with honeybees, is shown in the lower right-hand corner of the illustration. The two round windows in the fronts of these boxes are the critical areas in reference to which behavior of the bees is observed.

after completion of the transfer of the bees to the bait: each test provided four pairs of counts. The 32 pairs for the eight tests show 20 counts in which the number of bees at *EW* exceeded the number at *CW*, and 12 counts in which the count at *CW* equalled or was larger than for *EW*.

At the bottom of table 1 the totals for the populations found at *CW* and at *EW*, respectively, and for the four time intervals specified have been set down. In the eight tests a total of 133 bees were set at work in front of *CW* and 132 at *EW*. Each of these values was taken as 100 per cent and



on this basis the population percentages remaining at each of the windows after 5, 10, 15 and 20 minutes were computed. In all four pairs of values the percentage retained at *EW* exceeds that for *CW*. The per cent differences are +13.6, +21.6, +7.7, and +6.9, for 5, 10, 15 and 20 minute intervals, respectively, in favor of the experimental window. In summary this

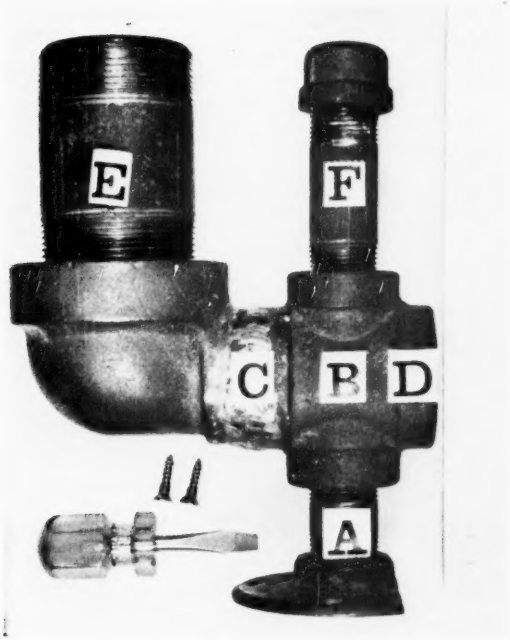


FIGURE 3

Lateral view of osmic radiation receiver which has been formed from black iron pipe fittings. *A*, receptacle for liquid; *B*, gas chamber above surface of liquid; *C*, closed back of chamber formed of black iron diaphragm; *D*, window sealing the front opening of the chamber; *E*, accessory chamber, open at top, for dry ice or other means of temperature control; *F*, an extension of air-tight chamber, *B*, filled with calcium chloride for absorbing water vapor. The unit as illustrated weighs 7 pounds. Scale in cm.

FIGURE 2

The upper view represents a pair of osmic radiation receivers mounted on a tripod-supported platform. The window in the right-hand unit will pass infra-red radiation into the air-tight gas chamber behind it, and is designated the "experimental" window. The control or blank unit is on the left in this view. The window in this unit will not pass infrared radiation into its gas chamber for the reason that a rear segment, made of ordinary glass, reflects such radiation.

The lower view represents the pair of units with the fronts of the housings removed. Plastic tubes 9 cm. in length and 5.5 cm. in diameter are shown fitted to the window collars and projecting horizontally. On the inside and at the bottom of each tube an attached bit of cloth provided a surface on which a small measured amount of honey bait was placed. In this view a number of bees may be seen working in each tube. Ordinarily the fronts of the housing boxes are in place.

TABLE 1

GROUPED DATA FOR EIGHT SUMMERTIME FIELD TESTS OF OLFACTORY RESPONSE IN HONEYBEES

TEST NO.	DATE, 1948	HOUR	BEE POPULATION AT START OF TEST		NUMBER OF BEES COUNTED IN BAIT TUBE AFTER 5 MIN.		AFTER 10 MIN.		AFTER 15 MIN.		AFTER 20 MIN.	
			CW	*EW*	*CW*	*EW*	*CW*	*EW*	*CW*	*EW*	*CW*	*EW*
1	8/17	5:06 P. M.	18	18	16	18	8	14	5	11	2	9
2	8/18	9:40 A. M.	22	22	10	22	6	16	5	3	2	0*
3	8/18	10:05 A. M.	18	18	16	18	11	15	7	10	4	2*
4	8/18	11:41 A. M.	18	18	18	18*	7	10	4	7	1	3
5	8/19	10:35 A. M.	18	18	16	18	11	9*	7	2*	0	2
6	8/19	12:20 P. M.	11	10	10	8*	10	7*	3	3*	0	3
7	8/19	2:23 P. M.	12	12	10	11	7	11	0	7	0	4
8	8/19	2:55 P. M.	16	16	10	10*	6	12	7	5*	8	3*
Totals: Pop. honeybees			133	132	106	123	66	94	38	48	17	26
Per cent of pop. after various intervals			100	100	79.6	93.2	49.6	71.2	28.6	36.3	12.8	19.7

* Pairs of populations in which *CW* (control window) is as large or larger than *EW* (experimental window).

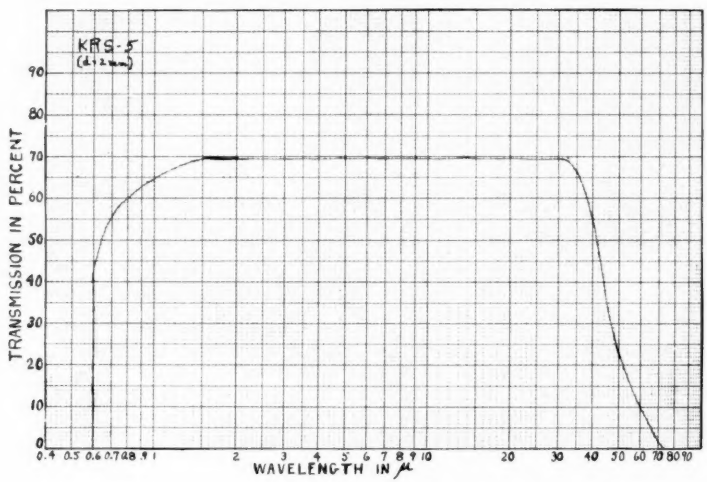


FIGURE 4

Transmission curve for thallium bromide iodide crystal window, 2 mm. in thickness.

group of eight tests, made under conditions when the bees were very actively working at near-by fields of bloom and therefore not easily distracted by artificial stimuli, yields results which tend to support the olfactory radiation hypothesis.

In the foregoing experiments the honey bait was placed inside plastic tubes which extended from the windows. The tubes were an easy fit on

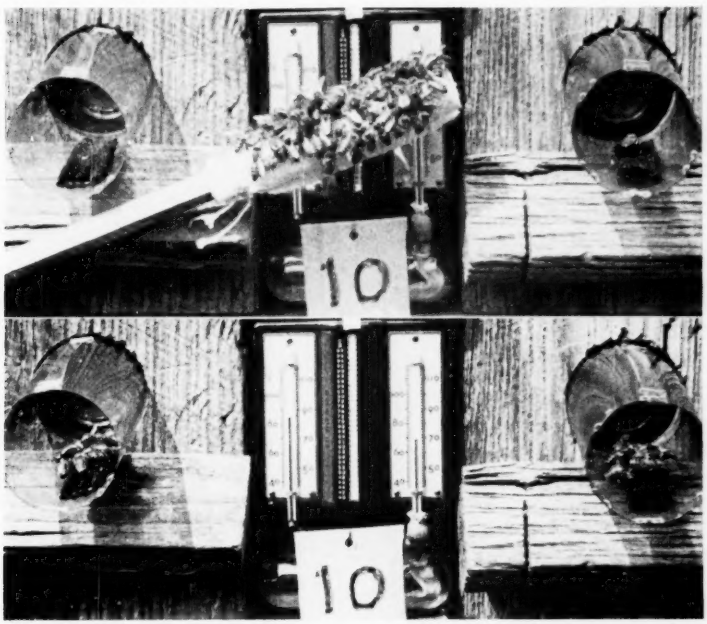


FIGURE 5

Transfer of honeybees to baited areas extending from the windows of the osmic radiation receivers. The upper view shows the transfer wand held in front of the units just before walking the bees off into the tubes. The lower view represents the condition immediately after the bees have been transferred to the baited tubes.

the iron collars surrounding the windows and also in the openings in the wooden housings. The tubes were completely open at their front ends. Tubes rather than flat projections were selected as supports for the bait because they would tend to hold the bait odor, locate the insects present in a favorable position for radiating into the windows, and restrain them from crawling about on the wood surrounding the windows. Furthermore, the tubes were easy to manipulate in placement, withdrawal and exchange.

FIGURE 6

A succession of photographic views which from top to bottom constitute the record of one experimental test. *A* shows the bees immediately after transfer to the bait areas; *B* represents the condition when the bait is about exhausted. Between *B* and *C* the bait tubes have been gently removed and the bees shaken out of them near the ground directly below the two units. *C* and the other views, to the bottom of the figure, illustrate at succeeding intervals of time the predominating tendency of the bees to gather on or near the experimental window, which is on the left in this illustration.

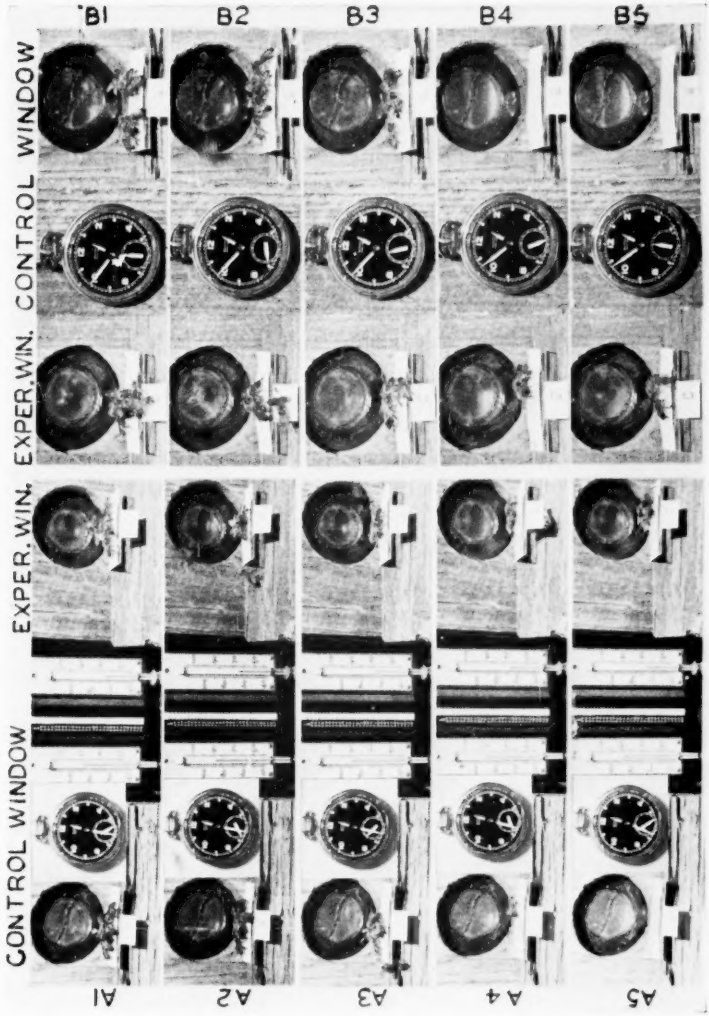


FIGURE 7

Two sets of photographic records representing two succeeding tests in which the experimental window is at the right in the first, and at the left in the second test. In place of tubes for holding the bait, small platforms have been arranged. For purposes of this illustration the photographs of the timepiece which during the experimentation were above the thermometers and near the tops of the unit housings, have been mounted in line with the windows to facilitate compactness of illustration. For discussion of the results of these two tests see the text.

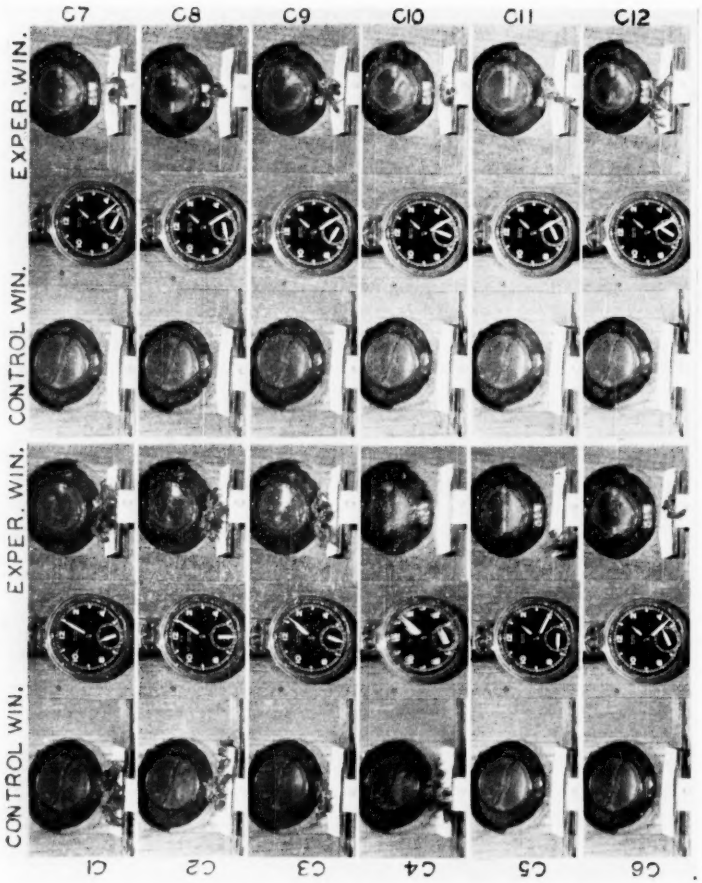


FIGURE 8

The 12 individual photographs composing this illustration constitute the record for one osmic radiation test. The beginning of the test is represented by the extreme upper left-hand view, C1, the end of the test by the bottom right-hand view, C12. When the test began platform *C* was below the experimental window. Immediately after the third view in the sequence, taken at 12:06, the two platforms were exchanged, as far as possible without disturbing the bees.

However, they have one disadvantage for the experimenter, since he finds them something of a visual obstruction, when attempting to count the bees present, and also inconvenient when attempting to record the populations photographically.

Early in November, 1948, a few experiments were tried using flat, plastic platforms attached at the bottom of each window collar and extending to the front and horizontally from that position. These platforms were two inches wide and one inch deep. By means of adhesive tape, a small amount of white cotton gauze was attached and covered the mid-portion of each portion. A series of three tests of this type are here reported. These were made at the same Connecticut farm at which our work had been done the previous fall.[10] In these tests a special effort was made to secure good photographic recording of the populations present at varying time intervals.[11] The season of the year was favorable; wild nectar was scarce. But we predicted that with the flat open platforms the bees would remain around the windows for a shorter time than when bait was placed in the tubular supports.

Test A was made shortly after 12:30 P.M.; the results are photographically illustrated in the left-hand portion of figure 7. In preparing this figure the original photographs have been cut down so as to include principally the two windows. The photos of the face of the timepiece have been brought down from their original position near the tops of the housings, and pasted between the windows for the sake of compactness in the illustration. The two platforms in these tests were lettered *W* for the control window and *C* (crystal) for the experimental window. The control window (*W*) was on the left, the experimental window on the right. Three drops of honey were carefully placed near the center of each platform, and eight worker bees were transferred to each platform. Photo A1 was taken immediately after the bees had been transferred and had started working. Fifteen seconds later, A2, eight bees still working in front of each window. Photo A3 made 26 seconds after A1, still shows most of the bees present on each side. In photo A4, made one minute and 16 seconds following A1, a single bee is on *W*, in front of the control unit, while three bees are on *C*, in front of the experimental window and two more are near about. The final photo, A5, taken two minutes after A1, reveals a characteristic result, there are no bees on platform *W* while five are shown on platform *C*.

Test B followed shortly after Test A and the five photographic records of it are reproduced in the right-hand portion of figure 7. It will be noted that the right-left positions of the two units have been exchanged, and in Test B the experimental window is now on the left-hand side. Again three drops of honey were placed centrally on each platform and eight worker bees were transferred to each platform. Photo B1 was made just

following this transfer. Photo B2 was made 43 seconds later. The bees had practically exhausted the available bait, and were now scattered and moving about. B3, taken 57 seconds after B1, indicates that some of the insects had left each platform. In Photo B4, recorded two minutes and 26 seconds after B1, four insects may be seen on platform *C*, near the experimental window, none on platform *W*. The final photo, B5, made three minutes after the beginning, records three bees on platform *C* and none on *W*. The reader will note that both of these tests, A and B, followed method (*b*) described as "the bait exhaustion method."

Test C, after a short interval, followed the previous two tests, and was made according to method (*c*) in which the bait platforms were exchanged in the early portion of the test. This experiment extended for 21 minutes, 12 photographic records were taken and all have been reproduced in figure 8. The two units had again been exchanged in their right-left position. Three drops of honey were placed on each platform and ten worker bees were transferred to each. Photo C1 records the conditions present immediately after the two equal populations had been set to work. It should be stated that two or three bees were transferred at a time to each platform, by moving the wand first to one and then to the other, in turn, so that neither platform would suffer the handicap that the bees got to work on it earlier than on the other one. In Photo C2, 34 seconds after C1, the bees on platform *W* had begun to move about and one or two had left. In Photo C3, 39 seconds after C1, nine or ten bees may be seen on platform *C*, only three on platform *W*. Shortly after this view was made no bees remained on platform *W*. Then the experimenter exchanged the positions of the two platforms without greatly disturbing the bees resting on platform *C*. As he completed the exchange, he unfortunately jostled the tripod of the camera so that the picture taken, Photo C4, was somewhat blurred. Nevertheless, it indicates that the bees were successfully transferred on platform *C* to the position in front of the control window. The bees soon began leaving platform *C;* this is not recorded in the picture sequence. For about 13 minutes no more than one bee at a time would appear on platform *W*. If another came, he would be fought off. Some would fly about the experimental window briefly. Photo C5, made nearly 13 minutes after Photo C4, and 20 minutes after C1, records two insects at platform *W*, and none on platform *C*. In fact, from 1:20 P.M., C5, until the end of the test, 1:26 P.M., C12, no bee alighted on platform *C* whereas on *W* by the experimental window the population ranged from two to five. The bees on *W* tended to fight each other and to fight off other bees that attempted to join them in their frustrated search for the honey which they seemed to sense was present.

Discussion.—The experiments on honeybees reported in this paper have constituted an attempt on our part to put the olfactory radiation

hypothesis to test. These field trials followed some laboratory tests on cockroaches carried out with an apparatus built according to the same principles embodied in the osmic radiation receivers here described. Our initial trial with bees was on September 1, 1947, at an apiary on a farm in Madison, Connecticut. In this test the beekeeper kindly coöperated with us. He introduced a small frame smeared with honey into the top of one of the hives. The smell of honey within the hive caused the bees to become active, in a search for honey, and within 15 minutes there were large numbers of them outside the hive on some frames from which the honeycombs had been removed, but small bits naturally still remained. We now considered that the bees had been sufficiently stimulated to find honey in their near vicinity. At first we placed our two units on a suitable tripod support, facing the sun, and with their backs about 20 feet away from the hive. Watching the units for 30 minutes, no bees were observed to pay any attention to either of the windows. We then changed the position of the units so as to face the hive and moved them about seven feet away from it. Again observing for 30 minutes, we noted that no bees came to either window. We then placed the units on the ground directly in front of the hive entrance and three feet away from it. The dry ice in the cooling chambers of each box had caused the iron window frames to frost up and moisture to gather on the windows. The moisture was removed and we observed the units for 15 minutes; no bees seemed to pay any attention to either box or window. The units were moved still nearer to the hive, and again no positive response. It was now 5:30 P.M.; we moved the boxes to within nine or ten inches from the mouth of the hive. Boards immediately below and in front of the hive formed a suitable platform for the units. There was activity at the mouth of the hive, workers were coming and going, also some workers were stinging drones and pushing them to one side. The two units were about one inch apart, and blocked low-angle approach of the bees to the hive opening. Under these conditions and after 15 minutes of watching, we observed no approach by the bees to the boxes themselves, that is, no bee lighted on either box, but one or two bees did pause a moment in flight in front of the experimental window. At this point we placed a thin string around both boxes in such a way that its position horizontally bisected each window. By means of a penknife blade, a small drop of honey was placed on the string directly in front of each window. After about five minutes there appeared a fly, then a yellowjacket. The yellowjacket perched on the string in front of the experimental window and consumed the drop of honey, then explored up and down the string immediately in front of the window. Presently he found the honey in front of the control window, where the fly had been engaged. The fly was frightened off and the yellowjacket consumed that drop of honey also. The yellowjacket now returned to the experimental window, re-

examined the string where the honey had been, then flew up in front of the window. He returned to the string in front of the control window, examined the string, but did not pay any obvious attention to the window. During the next 15 minutes a yellowjacket (apparently the same one) returned frequently, paused in front of the experimental window or lighted on the string or window or on the frame close to it. So far as we could judge, he paid about the same amount of attention to the string in front of the window as he did to the window itself. His behavior in reference to the experimental window was much more positive than in reference to the control window, in front of which he had gotten an equal amount of honey, and was last rewarded. These initial trials on our first day of experimenting led us to devise the technique of supplementary baiting in short plastic tubes projecting from the window areas.

It seems probable that the bait-tube technique can be improved. The flat platforms were considerably narrower than the length of the tubes. The front edges of the platforms extended only $1^{11}/_{16}$ inches or 43 mm. from the vertical planes of the front surfaces of the thallium windows. Tubes of this effective length or even shorter might be used and the bait surface disposed around an arc of the inside circumference rather than lengthwise in the tube. Transferred bees would then tend to have an initial orientation facing inward toward the bait and also toward the windows. Such refinements together with a technique for marking the bees set to work at the window areas might yield more quantitative and more analytical results than we have thus far secured.

In the tests reported above we have varied the position and distance of the receiver units from the hive. The two units alternated in their respective right-left positions. The temperatures of the two units were kept nearly equal, and below the environmental temperature but not always at the same differential between environment and the units. Different times of day in mid-summer and in late fall have been tried. The bait tubes were used interchangeably, that is, one of them was not made to be used at the experimental window, and the other at the control window. The tubes were sometimes (method *c*) exchanged after the beginning of a test, sometimes in other tests (method *a*) they were removed after the experiment was well started. The windows were often wiped dry if they tended to gather moisture, and they were kept clean. Honey from the same colony worked with was always used as the bait. Throughout all these controls and variations of testing the bait-seeking behavior of bees introduced into the situation seems much more positive in reference to the infrared passing window. It seems fair to assume that in front and near this window the insects have an olfactory experience and that it is this which causes them to collect and to remain for some minutes. After the bait has become exhausted the situation, for those that remain, would seem to offer no primary

reward. These bees sensing honey, perhaps strongly, but not finding it seem irritable. At least it is characteristic of their behavior that in the latter half of a test they fight off other comers that approach the experimental window. Except for this behavioral factor the population differential between the windows might have been more striking. A technique of working with a group of bees, one at a time, as reported by Frings,[12] should avoid any such masking effect introduced by this element of social interaction. Our field experiments unfortunately lacked this type of control.

If it is fair to assume that the bees collect at *EW* because of an olfactory experience gotten there which is not found at *CW*, then on the physical side it may be assumed that the behavioral difference is based on a difference in infrared radiation. The experiments indicate that the difference is due to radiation lost by the insects, in all probability from their antennae.[2, 12, 13] The experiments reported here do not settle the question whether the radiation is direct to the cooled gas molecules of honey vapor in the air-tight chamber beyond the thallium window, or whether ambient honey-vapor molecules adsorbed on the antennae provide a filter-like film through which radiation takes place, and is accentuated by the cooled chamber behind the transmitting window near the insects. In the latter case it would probably make little difference whether the chamber was filled with honey gas or some other gas. On the basis of some other experimentation not here reported, we are inclined to believe that the second alternative is the one that is operating. In either case a radiation hypothesis, as the stimulus for olfaction in honeybees, seems to be supported.

Conclusions.—1. It is suggested that the initial physical event in the stimulation of the olfactory receptors is due to an increase in their heat radiation caused by the heat absorption of certain vapors or gases.

2. Osmic radiation receivers, designed to provide infrared passing windows for air-tight gas chambers, are described.

3. A technique and methodology of using osmic radiation receivers in field studies on honeybees is outlined.

4. Experimental tests on the behavior of honeybees in reference to the osmic radiation receivers are described and critically evaluated. Controls were introduced for a number of variables.

5. In general, the results support the hypothesis. More bees definitely paid attention and attended for longer intervals to the infrared passing window than to the window of the control unit. The experiments support the interpretation that the behavioral difference is due to increased heat losses from the insects' antennae.

* This investigation was supported in part by grants from Fluid Research Funds, Yale University, and from the Research Corporation of New York City.

[1] Beck, L. H., and Miles, W. R., "Some Theoretical and Experimental Relationships between Infrared Absorption and Olfaction," *Science*, **106**, (A) 511 (1947).

[2] Wigglesworth, V. B., *The Principles of Insect Physiology*, E. P. Dutton & Co., New York (1939), pp. 147–152. Snodgrass, R. E., "The Morphology of Insect Organs and the Sensory Nervous Systems," *Smithsonian Inst. Pub. Misc. Collections*, **77** (8), 1 (1926).

[3] McCord, C. P., and Witheridge, W. N., *Odors: Physiology and Control*, McGraw-Hill Book Co., New York (1949). See extensive bibliography.

[4] Three brief abstracts concerning KRS-5 (composed of mixed crystals of thallium bromide and thallium iodide) appear in *J. Opt. Soc. Am.*, **38**, 663–664 (1948).

[5] The KRS-5 discs were produced at Engineer Research and Development Laboratories, The Engineer Center, Fort Belvoir, Va. C. R. Hoover, Project Engineer, Radiation Engineering Branch, supplied the transmission chart for KRS-5 (thickness of 2.0 mm.) and stated in a letter dated July 3, 1947, that the "transmission characteristics of KRS-5 are no longer classified. . . ." Paul J. Ovrebo of Wright Field, through his kindness, was instrumental in helping us secure the use of two discs of this recently perfected and expensive material.

[6] Product of Lakeside Chemical Co., So. Chicago Ave., Chicago. Mr. Beardsley of Wright Field kindly gave us the benefit of his experience in the use of this material.

[7] The KRS-5 disc in the control unit developed an observable crack while being sealed in its collar against the glass. We had no alternate disc to substitute and therefore sealed the crack. This seal across the face of the disc may be noted in some of the reproduced photographs.

[8] Cooke, F., "Optical Finishing of KRS Materials," *J. Opt. Soc. Am.*, **37**, (A) 948 (1947). High polish can be obtained through the use of pitch and beeswax laps.

[9] We wish to acknowledge the kind coöperation of Frederick A. Godley, owner and operator of this farm in New York.

[10] We wish to acknowledge the kind coöperation of William C. Peck, owner and operator of this Connecticut farm where we have conducted most of our tests on honeybees.

[11] *Life* Magazine photographer Eric Schaal was happily with us on this occasion. We have permission to use some of the photographs which he took of our experiments. See figures 7 and 8.

[12] Frings, H., "The Loci of Olfactory End-Organs in the Honey-Bee, *Apis mellifera* Linn," *J. Exper. Zool.*, **97**, 123–134 (1944).

[13] Crozier, W. J., "Chemoreception," Ch. 19, pp. 987–1036, *Handbook of General Experimental Psychology*, Murchison, C., Editor, Clark Univ. Press, Worcester, Mass. (1934).

[14] Beck, L. H., "Osmics: Theory and Problems Related to the Initial Events in Olfaction," *Medical Physics*, Glasser, Otto, Editor. Second edition in press.

THE BEADED MINUTE COMBINATION AND SEX-DETERMINATION IN DROSOPHILA

BY A. H. STURTEVANT

KERCKHOFF BIOLOGICAL LABORATORIES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

Communicated March 19, 1949

In a recent paper[1] Goldschmidt reports that Beaded Minute males of *Drosophila melanogaster* are usually intersexual, the effect being produced by any one of 10 out of 12 different Minutes tested. From experiments with these types he draws certain conclusions concerning the mechanism of sex-determination in Drosophila.

I have examined Beaded Minute flies in which eight different Minutes were concerned (including three of those from which Goldschmidt reports more extreme effects), and have found numerous aberrant males in seven of these combinations; but detailed study has convinced me that these are not intersexes. They represent cases in which the external genitalia are imperfectly developed or completely absent, but there is no clear indication of femaleness, either in the external genitalia or in any of the internal parts. The male genital ducts are often rudimentary, but neither they nor the gonads show any trace of femaleness. In normal males and females there is a sexual difference in the anal plates (two lateral ones in the male, a dorsal and a ventral in the female), and it has been shown[2, 3] that one of the most sensitive indices of intersexuality is the presence of three plates (two laterals and a ventral). I have carefully examined over 350 Beaded Minute males, from many cultures, with different Minutes and different genetic backgrounds, that were grown at widely different temperatures; in no case have I found any ventral anal plate. It is, however, frequently the case that the whole anal and genital area is smooth, with no sign of anal or genital opening or sclerite. Such individuals, when dissected, show very rudimentary genital ducts (usually not found at all, but perhaps overlooked), and a gut that is closed behind the rectal glands, its blind end lying free in the body cavity. These males, without exception, have had two normal sex-combs, and normal male-type abdominal tergites. The published accounts of triploid intersexes do not indicate any tendency to abortion of the anal plates, and I have recently examined over 200 triploid intersexes of varying grades without finding any case in which less than two anal plates were present.

The above account was submitted to Dr. Goldschmidt, who has kindly given me his comments on it. (See the paper immediately following this one.) He agrees with many of the observations, but has made additional ones which convince him that these specimens are really intersexes. Some

of these observations concern the shapes of the parts in the more extreme specimens, and leave me unconvinced. Dr. Goldschmidt has, however, called my attention to a few hairs that are often present on the seventh sternite (sixth in the terminology of Newby[3]) of Beaded Minute males, in positions very similar to those regularly found on the seventh sternite of normal females. I had overlooked these hairs, but can now confirm the observation that they are often present. However, I find that they are occasionally present in normal (and fertile) males of a wide variety of constitutions. They were even present in an XY male that was homozygous for the gene *tra*,[6] and that therefore can scarcely be considered as in any degree intersexual. These hairs are, then, not good indications of intersexuality. Their frequent presence in Beaded Minute males is interesting from a developmental point of view, but is not to be taken as a sign of femaleness.

We are then dealing not with intersexuality but with incomplete development of the anal and genital imaginal discs—a condition that is well known in a wide variety of Drosophila types as an occasional anomaly, and is a regular characteristic of certain genotypes—e.g., *In (2) ho* (Lewis[4]). In agreement with this conclusion is the fact that Beaded Minute *females* also often lack anal plates and external anal opening, though this happens less frequently than in Beaded Minute males. In five Beaded Minute females the genital opening was also absent. The tendency to imperfect development of imaginal discs, in Beaded Minute of both sexes, is most marked at the posterior end, but is still much greater than is usually the case in other parts of the body. Among the defects rather frequently seen may be mentioned roughened eyes with reduced area, mis-shapen legs, incompletely unfolded wings, and incompletely chitinized abdominal tergites.

Goldschmidt argues that the primary effect here is that of Beaded, the Minutes serving to accentuate it. My observations do not support this view, since I have seen such aberrant specimens in rather casual examinations of seven out of fifteen Minute strains with no Beaded ancestry. This result suggested that the effect is part of the "Minute reaction" studied by Schultz.[5] In accord with this conclusion is the fact that Hairless largely suppresses the effect, as it does the "Minute reaction." Delta accentuates the "Minute reaction," and accordingly Delta-3 was used. Delta-3 with Minute-3w or with Minute-S13 does in fact give a series of males and females showing the same anal and genital defects as does Beaded Minute; but so also does Delta-3 Beaded, and I have found an exactly similar series in a wholly unrelated culture in which only a few second chromosome mutant genes (not Minutes) were involved. In all these cases the series of types found is very closely similar to those in Beaded Minute, and the frequencies are well within the range of variation among strongly affected Beaded Minute cultures. It is clear, therefore, that this type of

defect is rather easily produced, by a wide variety of genetic combinations.

Goldschmidt suggests that the Beaded section of the third chromosome may turn out to be one of the two male determiners of Drosophila. This suggestion rests on the two conclusions that I have here suggested are incorrect—namely, that the Beaded Minute aberrant males are intersexual, and that their peculiarities are due primarily to the Beaded locus.

Notes on Dr. Goldschmidt's Paper That Immediately Follows This One.—On many points Goldschmidt and I are in agreement. It seems clear that the reduction of the anal plates in both sexes is not in itself relevant to the question of sex determination. I should also conclude that reduction of the male genital apparatus, and the appearance of hairs on the seventh sternite, are indicative of intersexuality only if they can be shown to be associated with other, more unambiguous, female characteristics. In my material they were not so associated; Goldschmidt finds them to be so, but only occasionally—as he suggests, perhaps owing to genetic differences between the strains used. Not having found Beaded Minute males with such properties, I can only conclude that they arise as a consequence of some other, unanalyzed, genetic composition—and that discussion of their relation to general theories of the genetic determination of sex is premature.

Summary.—Beaded Minute males are not intersexual, but are often aberrant because of incomplete development of the anal and genital imaginal discs—as are also Beaded Minute females.

Even if they are considered to be intersexual, the abnormality may be induced by any one of at least seven different Minutes, or by other independent combinations of genes, and therefore cannot be used to support the view that there are very few loci involved in the determination of sex.

[1] Goldschmidt, R. B., PROC. NAT. ACAD. SCI., **34**, 245–252 (1948).
[2] Dobzhansky, T., and Bridges, C. B., *Amer. Nat.*, **62**, 425–434 (1928).
[3] Newby, W. W., *Univ. Texas Pub.* 4228, 113–145 (1942).
[4] Lewis, E. B., *Genetics*, **30**, 137–166 (1945).
[5] Schultz, J., *Ibid.*, **14**, 366–419 (1929).
[6] Sturtevant, A. H., *Ibid.*, **30**, 297–299 (1945).

THE INTERSEXUAL MALES OF THE BEADED MINUTE COMBINATION IN DROSOPHILA MELANOGASTER

BY RICHARD B. GOLDSCHMIDT

UNIVERSITY OF CALIFORNIA

Communicated March 26, 1949

Dr. A. H. Sturtevant kindly sent me a copy of the manuscript of his paper on the Beaded Minute males of *Drosophila melanogaster*. He holds that these males described by me as intersexual males are not real intersexes, but products of rudimentation of the genital and anal discs. There is no doubt that many individuals might be described this way if they stood alone. But this is also true for many triploid intersexes of Dobzhansky's classes I–III (Dobzhansky, 1930). Actually, these types are only one in a series of morphological changes which indicate clearly that we are dealing with male intersexes. It seems that, by chance, Sturtevant met only with the one type which usually, though not always, is the most frequent one. Detailed statistics on the frequency and genetics of the different types will be published in time together with the necessary illustrations of all the types, grades and transitions.

It should be stated first that the lower grade Bd-M intersexes are different from the triploid intersexes (Dobzhansky's classes I, II). Both the anal plates and the type of reduction of the armature are very different. Sturtevant is much impressed by this fact. But it should be pointed out that, thus far, the only male intersexes available for comparison are the triploid ones. For female intersexes a number of different strains are known, all of which have their specific features, which show that the details of intersexual transformation vary with the different genetic backgrounds. Thus, I cannot regard the triploid intersexes as an absolute standard of intersexual structure. But it is a fact that the internal structure of the Bd-M intersexes is exactly the same as that described by Dobzhansky for the triploids of his classes I–III. Unfortunately, the higher classes of intersexuality are rather rare in the Bd-M case.

As Sturtevant mentions only the one type of intersexuality which I call the "reduction type" it should be proved first that this type is also actually an expression of intersexuality in which the reduction and final disappearance of the male genital armature is more prominent than the secondary appearance of female characters. This type is a combination of two phenomena, namely, the reduction and disappearance of the anal plates and that of the genital armature proper. It is the conspicuous reduction of the anal plates found in this type only, but absent in the other types, which led Sturtevant astray. I think that his statements and interpretations are correct as far as the anal plates are concerned. In a rather good correlation with the intensity of the Bd-M effect upon the males, the anal plates and even the

abdominal tip are reduced in a part of the females, and some not Beaded-Minutes may show this also. But it will be shown below that the involution of the anal plates is independent of the reduction of the genital armature. In the reduction type, now under discussion, the genital armature is retracted and found (if not completely lost) below the surface, in conditions which indicate that it had developed to a certain degree—all degrees from completeness to total absence—and then stopped growing. (This includes also stoppage at different times of the well-known 360° rotation in the pupa.) These conditions parallel closely those described by Dobzhansky for the triploid intersexes, although not in every detail. The highest grade of this reduction type is characterized by a rather large abdomen without any genital structure at the tip. As obviously the entire disc region degenerated early, no further female differentiation is possible. But in less extreme grades (called in my work class III) the disc is still present below the integument, and in this stage I twice found (in not too many dissections of this specific type, as yet) individuals with rudimentary testes, without any male ducts, but with a complete vagina with its characteristic histological differentiation, and with a ventral receptacle complete in one, not perfect in the other case. In the higher grades of this type, hairs are found rather frequently on the 7th sternite. This is a female character. Examination of innumerable normal brothers of the same genetic constitution never showed this phenomenon.

A second type of intersex appears rather irregularly in the F_1 and F_2 broods. In our earlier work it was just as frequent as the reduction type in certain Bd-M combinations. Recently it was found rather rarely, indicating that specific modifiers may control these types. (All types may appear in the same brood. Statistics will be published in time.) In this second type there is no degeneration of the anal plates. Instead, they even hypertrophy. They grow out beyond the tip of the abdomen into two large plates which frequently unite dorsally (a female trait). The dorsal part of the genital arch to which the plates are attached is also pushed out as a socket, which, if sufficiently large, develops hairs of the type of the female 8th tergite, which it clearly represents. The genital armature proper undergoes the same stages of reduction as described for the first type until the ventral side of the abdomen below the large beak-like protruding anal plates is free of differentiations. In still higher grades this ventral surface becomes depressed and forms a deep cup. In more extreme individuals, the anterior lip of the cup-like depression grows backward until only a curved slit remains. In the only three dissections of this type made thus far the bottom of the covered cup was found to have the histological structure of the vagina (uterus) and is therefore considered as the secondary invagination of a vagina. (I agree with those authors who think the vaginal disc is not identical with the male disc.)

A third type, which we call the direct type is rather frequent in its lower grades, but thus far rather rare in the higher grades. It represents a simple and direct transformation of the male abdominal tip into an almost female one. There are, of course, also transitions and combinations between the three main types, e.g., degeneration of the anal plates with progressive differentiation of the female segment. The only male parts to become rudimentary in this type are the claspers and the penis, which assume very characteristic shapes during this process. The first step is the fusion of the two ventral tips of the genital arch with their characteristic tufts of hair into a single median one, which protrudes as a knob. The genital arch then broadens and becomes oriented as an 8th segment. The possibility that a 9th segment is involved will be discussed another time. The knob becomes larger and is transformed into a protruding structure—still with the tufts of hair—resembling the female abdominal tip and assumes the spatial relation to the anal plates typical of the female. Finally the genital arch forms a typical female eighth (or 8th and 9th?) segment with its normal pigmentation and arrangement of hairs. Also, the tuft of hairs on the "knob" scatters over the sternite. In one such extreme type the anterior surface of the knob carried a disc which might have been a vaginal disc or the anlage of the vaginal plates. Only rarely are the anal plates arranged in a dorsal and ventral position, as in females.

An accurate record of these facts requires many illustrations which will be published when all dissections are completed. It should be added that in these experiments thousands of brothers of the intersexes of other genetic constitutions, i.e., Bd or M or neither, were constantly watched as controls without a single case of intersexuality or reduction of the genital armature occurring. There is no doubt in my mind that the types described are male intersexes, which leaves the theoretical considerations which I made in my earlier paper (Goldschmidt, 1948) and in another one now in press (Goldschmidt, 1949) as they were originally formulated.

Dobzhansky, Th., *Bull. Bur. Genetics* (Leningrad), **8,** 91–158 (1930).
Goldschmidt, R. B., PROC. NAT. ACAD. SCI., Wash., **34,** 245–252 (1948).
Goldschmidt, R. B., *Arch. Julius Klaus-Stiftung*, **23,** 1948: 539–549.

A THEORY OF TRACE-ADMISSIBLE ALGEBRAS

BY A. A. ALBERT

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF CHICAGO

Communicated April 30, 1949

1. Introduction.—There have been a number[1] of studies of non-associative algebras in which the lack of the associative law has been compensated for by the assumption of one or more other identities. The principal tool in most of these studies has been a trace argument,[2] and the simple algebras, which are the end results, have invariably been of relatively special types.

In the present note we shall make a major advance in the theory of non-associative algebras by obtaining a general theory which includes all previous theories in which the radical was defined to be the maximal nilideal. We shall restrict our attention to power associative algebras over a field $\mathfrak{F}$ whose characteristic is prime to thirty, and shall call such and algebra $\mathfrak{A}$ *trace-admissible* if there is a bilinear function $\tau(x, y)$ with arguments x and y in $\mathfrak{A}$ and values in $\mathfrak{F}$ such that

$$\tau(x, y) = \tau(y, x), \qquad \tau(x, yz) = \tau(xy, z), \tag{1}$$

$\tau(e, e) \neq 0$ if e is an idempotent of $\mathfrak{A}$, $\tau(x, y) = 0$ if xy is a nilpotent element of $\mathfrak{A}$ or is zero. We define the *radical* of $\mathfrak{A}$ to be its maximal nilideal[3] $\mathfrak{N}$, and say that $\mathfrak{A}$ is *semisimple* if $\mathfrak{N} = 0$. We also say that $\mathfrak{A}$ is *simple* if $\mathfrak{A}$ is semisimple and $\mathfrak{A}$ and 0 are the only ideals of $\mathfrak{A}$.

If $\mathfrak{A}$ is any algebra there is an attached commutative algebra $\mathfrak{A}^{(+)}$. This is the same vector space as $\mathfrak{A}$ but is an algebra defined by a product $x \cdot y$ expressible in terms of the product xy of $\mathfrak{A}$ by the formula $2x \cdot y = xy + yx$. We shall show that when $\mathfrak{A}$ is trace-admissible then $\mathfrak{A}$ and $\mathfrak{A}^{(+)}$ have the same radical. Moreover, when $\mathfrak{A}$ is semisimple, the trace function $\tau(x, y)$ of $\mathfrak{A}$ is an admissible trace function for $\mathfrak{A}^{(+)}$. But then we can show that $\mathfrak{A}^{(+)}$ *is a Jordan algebra,* $\mathfrak{A}$ *is flexible,* $\mathfrak{A}$ *is simple if and only if* $\mathfrak{A}^{(+)}$ *is simple.* It follows that *the only simple trace-admissible algebras* are the (commutative) *Jordan algebras,* the *flexible* (noncommutative) *algebras of degree two,* and the (noncommutative flexible) *quasiassociative algebras.* These results show that unless some technique other than a trace argument is utilized in studying power associative algebras no new simple algebras will arise.

2. Elementary Properties.—If e is an idempotent of a power associative algebra $\mathfrak{A}$ over $\mathfrak{F}$ of characteristic prime to 30 it is known[4] that $\mathfrak{A}$ can be decomposed as the supplementary sum

$$\mathfrak{A} = \mathfrak{A}_e(1) + \mathfrak{A}_e(1/2) + \mathfrak{A}_e(0) \tag{2}$$

of subspaces $\mathfrak{A}_e(\lambda)$ such that

$$ea_\lambda + a_\lambda e = 2\lambda a_\lambda \tag{3}$$

for every a_λ of $\mathfrak{A}_e(\lambda)$. Then it is actually true that $ea_1 = a_1e = a_1$, $ea_0 = a_0e = 0$. The subspaces $\mathfrak{A}_e(1)$ and $\mathfrak{A}_e(0)$ are subalgebras of $\mathfrak{A}^{(+)}$, and this implies that $a_\lambda b_\lambda + b_\lambda a_\lambda = c_\lambda$ is in $\mathfrak{A}_e(\lambda)$ for every a_λ and b_λ of $\mathfrak{A}_e(\lambda)$ with $\lambda = 0, 1$. It is also true that

$$a_\lambda b_{1/2} + b_{1/2}a_\lambda = c_{1/2} + c_{1-\lambda} \tag{4}$$

where $c_{1/2}$ is in $\mathfrak{A}_e(1/2)$ and $c_{1-\lambda}$ is in $\mathfrak{A}_e(1 - \lambda)$ for every a_λ of $\mathfrak{A}_e(\lambda)$ and $b_{1/2}$ of $\mathfrak{A}_e(1/2)$ with $\lambda = 0, 1$. Finally

$$a_{1/2}b_{1/2} + b_{1/2}a_{1/2} = c_0 + c_1, \tag{5}$$

where the notations should be evident at this point.

The subspaces $\mathfrak{A}_e(1)$ and $\mathfrak{A}_e(0)$ are not, in general, subalgebras of $\mathfrak{A}$. However, since they are subalgebras of $\mathfrak{A}^{(+)}$, and powers in $\mathfrak{A}$ and in $\mathfrak{A}^{(+)}$ coincide, it is true that the polynomial algebra $\mathfrak{F}[a_\lambda]$ is contained in $\mathfrak{A}_e(\lambda)$ for every a_λ of $\mathfrak{A}_e(\lambda)$, $\lambda = 0.1$.

The results about the decomposition above are consequences, in the case where $\mathfrak{A}$ is commutative, of the identity $x^2x^2 = (x^2x)x$. If we replace x by $x + \phi y$ in this identity the coefficient of ϕ must be zero and this yields the basic identity[5]

$$yx^3 = 4(yx)x^2 - (yx^2)x - 2[(yx)x]x. \tag{6}$$

We shall use this identity in the proof of an essential lemma.

LEMMA 1. *Let $\mathfrak{B}$ be an ideal of a power associative algebra $\mathfrak{A}$ and e be an idempotent of $\mathfrak{A}$ such that $\mathfrak{A}_e(0)$ is contained in $\mathfrak{B}$. Then g^3 is in $\mathfrak{B}$ for every g of $\mathfrak{A}_e(1/2)$.*

For $\mathfrak{B}$ is also an ideal of $\mathfrak{A}^{(+)}$, and g^3 is the same in both $\mathfrak{A}$ and $\mathfrak{A}^{(+)}$. We may then assume, without loss of generality, that $\mathfrak{A}$ is commutative. We use (5) to write $g^2 = c_0 + c_1$ and put $x = g$, $y = e$ in (6) to obtain $e[(c_0 + c_1)g] = 2g^3 - c_1g - g^3 = (c_0 + c_1)g - c_1g = c_0g$. Since c_0 is in the ideal $\mathfrak{B}$ so are $e(c_0g)$ and c_0g. Then $e(c_1g)$ is in $\mathfrak{B}$. But $c_1g = d_{1/2} + d_0$ by (4), $e(c_1g) = {}^1/_2 d_{1/2}$ is in $\mathfrak{B}$, d_0 is in $\mathfrak{B}$, c_1g is in $\mathfrak{B}$, $(c_1 + c_0)\,g = g^3$ is in $\mathfrak{B}$ as desired.

3. Trace Admissible Algebras.—Let $\mathfrak{A}$ be power associative, e be an idempotent of $\mathfrak{A}$, $\tau(x, y)$ be an admissible trace function for $\mathfrak{A}$. Use (2) and note that

$$\tau(x_1, y_0) = \tau(x_1, y_{1/2}) = \tau(x_0, y_{1/2}) = 0 \tag{7}$$

for every x_λ of $\mathfrak{A}_e(\lambda)$ and y_μ of $\mathfrak{A}_e(\mu)$. Indeed $x_1y_0 = 0$, $\tau(x_\lambda, y_{1/2}) = \tau(x_\lambda, ey_{1/2} + y_{1/2}e) = \tau(ex_\lambda + x_\lambda e, y_{1/2}) = 2\lambda\tau(x_\lambda, y_{1/2})$ and so, if $2\lambda \neq 1$, we have $\tau(x_\lambda, y_{1/2}) = 0$. Observe that we have used the consequence

$$\tau(x, yz) = \tau(zx, y) \tag{8}$$

of (1).

Any admissible trace function for $\mathfrak{A}$ is trivially an admissible trace function for any subalgebra $\mathfrak{B}$ of $\mathfrak{A}$. It is known[6] that the *radical* (maximal ideal of nilpotent elements) of $\mathfrak{A}$ *is the set* $\mathfrak{N}$ *of all elements* x *of* $\mathfrak{A}$ *such that* $\tau(x, y) = 0$ *for every* y *of* $\mathfrak{A}$. Thus $\mathfrak{N}$ is independent of τ.

An idempotent e of $\mathfrak{A}$ is said to be *principal* if there exists no idempotent u orthogonal to e. Since every element of $\mathfrak{A}_e(0)$ is orthogonal to e the space $\mathfrak{A}_e(0)$ contains no idempotent. But if a were a non-nilpotent element of $\mathfrak{A}_e(0)$ the algebra $\mathfrak{F}[a]$ is associative and would contain an idempotent w which would be in $\mathfrak{A}_e(0)$. Hence all elements of $\mathfrak{A}_e(0)$ are nilpotent when e is principal. We may then prove[7]

LEMMA 2. *Let* e *be a principal idempotent of a trace admissible algebra* $\mathfrak{A}$ *with radical* $\mathfrak{N}$. *Then* $\mathfrak{A}_e(1/2) + \mathfrak{A}_e(0)$ *is a subspace of* $\mathfrak{N}$.

For the sum $x_0 + y_0$ of any x_0 and y_0 of $\mathfrak{A}_e(0)$ is nilpotent, so is $(x_0 + y_0)^2$, $0 = \tau(x_0 + y_0, x_0 + y_0) = \tau(x_0, x_0) + \tau(y_0, y_0) + 2\tau(x_0, y_0)$. But x_0^2 and y_0^2 are nilpotent, $\tau(x_0, x_0) = \tau(y_0, y_0) = 0$, $\tau(x_0, y_0) = 0$. By (7) $\tau(x_0, y) = 0$ for every y of $\mathfrak{A}$, $\mathfrak{A}_e(0)$ is a subspace of $\mathfrak{N}$. By Lemma 1 we see that g^3 is in $\mathfrak{N}$ for every g of $\mathfrak{A}_e(1/2)$, g is nilpotent. Then $\tau(x_{1/2} + y_{1/2}, x_{1/2} + y_{1/2}) = 2\tau(x_{1/2}, y_{1/2}) + \tau(x_{1/2}, x_{1/2}) + \tau(y_{1/2}, y_{1/2}) = 0$, $\tau(x_{1/2}, y_{1/2}) = 0$ for every $x_{1/2}$ and $y_{1/2}$ of $\mathfrak{A}_e(1/2)$, $\tau(x_{1/2}, y) = 0$ for every y of $\mathfrak{A}$ by (7), $\mathfrak{A}_e(1/2)$ is contained in $\mathfrak{N}$.

4. Reduction to the Semisimple Case.—If $\mathfrak{A}$ has a unity quantity 1 we may define $\tau(x) = \tau(1, x)$ so that $\tau(x)$ is a *linear* function of x. Then $\tau(x, y) = \tau(1, xy) = \tau(xy)$ by (1) and the bilinear function $\tau(x, y)$ is the usual linear trace function τ of a product. Note that $\tau(e, e) = \tau(e) \neq 0$ for any idempotent e, $\tau(y) = \tau(1, y) = 0$ for y any zero or nilpotent quantity, $\tau(xy) = \tau(yx)$, $\tau[x(yz)] = \tau[(xy)z]$. We now prove

LEMMA 3. *If* $\tau(x, y)$ *is an admissible trace function for an algebra* $\mathfrak{A}$ *with a unity quantity then* $\tau(x, y)$ *is an admissible trace function for* $\mathfrak{A}^{(+)}$ *so that the radicals of* $\mathfrak{A}$ *and of* $\mathfrak{A}^{(+)}$ *coincide.*

For $\tau(x, y) = \tau(y, x)$ is a property only of the vector space $\mathfrak{A}$, $\tau(x, y \cdot z) = {}^1/_2\tau(x, yz + zy) = {}^1/_2\tau(xy + yx, z) = \tau(x \cdot y, z)$ by (8) and (1). If e is an idempotent of $\mathfrak{A}^{(+)}$ then e is also an idempotent of $\mathfrak{A}$ and so $\tau(e, e) \neq 0$. Assume now that $x \cdot y = {}^1/_2(xy + yx)$ is nilpotent. Then $\tau(x \cdot y) = 0$ as was shown above, ${}^1/_2\tau(xy + yx) = \tau(xy) = \tau(x, y) = 0$ as desired. Note that actually $\tau(x \cdot y) = \tau(xy)$ for every x and y of $\mathfrak{A}$.

We shall next prove

THEOREM 1. *Let* $\mathfrak{A}$ *be a trace-admissible power associative algebra over a field* $\mathfrak{F}$ *of characteristic prime to 30, and* $\mathfrak{N}$ *be the radical of* $\mathfrak{A}$. *Then the semisimple algebra* $\mathfrak{A} - \mathfrak{N}$ *is trace-admissible and has a unity quantity so that* $(\mathfrak{A} - \mathfrak{N})^{(+)} = \mathfrak{A}^{(+)} - \mathfrak{N}^{(+)}$ *is semisimple,* $\mathfrak{N}^{(+)}$ *is the radical of* $\mathfrak{A}^{(+)}$.

For the elements of $\mathfrak{A} - \mathfrak{N}$ are the images x_0 of the elements x of $\mathfrak{A}$ under a homomorphism $x \to x_0$ of $\mathfrak{A}$ onto $\mathfrak{A} - \mathfrak{N}$. Define $\tau(x_0, y_0) = \tau(x, y)$. Then if also $x_1 \to x_0$ and $y_1 \to y_0$ we know that $x_1 = x + x'$, $y_1 = y + y'$ where x' and y' are in $\mathfrak{N}$, $\tau(x_1, y_1) = \tau(x, y) + \tau(x', y) + \tau(x, y') + \tau(x', y') = \tau(x, y)$ since $x'y$, xy' and $x'y'$ are in $\mathfrak{N}$. The properties (1) are trivial for the function $\tau(x_0, y_0)$ since they hold for $\tau(x, y)$, and if x_0y_0 is nilpotent then a power of xy must be in $\mathfrak{N}$, xy is nilpotent, $\tau(x, y) = \tau(x_0, y_0) = 0$. Every idempotent e_0 of $\mathfrak{A} - \mathfrak{N}$, in the power associative case, is the image of an idempotent e of $\mathfrak{A}$ and $\tau(e_0, e_0) = \tau(e, e) \neq 0$, $\tau(x_0, y_0)$ is an admissible trace function for $\mathfrak{A} - \mathfrak{N}$.

The algebra $\mathfrak{A} - \mathfrak{N}$ is necessarily semisimple and is now trace-admissible. Let e be a principal idempotent of $\mathfrak{C} = \mathfrak{A} - \mathfrak{N}$. By Lemma 2 we see that $\mathfrak{C}_e(1/2) + \mathfrak{C}_e(0) = 0$, e is the unity quantity of $\mathfrak{C}$. By Lemma 3 the algebra $\mathfrak{C}^{(+)}$ is semisimple. But every ideal $\mathfrak{B}$ of $\mathfrak{A}$ determines an ideal $\mathfrak{B}^{(+)}$ of $\mathfrak{A}^{(+)}$, $\mathfrak{A} - \mathfrak{B}$ and $\mathfrak{A}^{(+)} - \mathfrak{B}^{(+)}$ are the same vector spaces. Hence $(\mathfrak{A} - \mathfrak{N})^{(+)} = \mathfrak{A}^{(+)} - \mathfrak{N}^{(+)}$, $\mathfrak{N}^{(+)}$ must be the radical of $\mathfrak{A}^{(+)}$.

5. *The Principal Results.*—The results just derived show that *if $\mathfrak{A}$ is a semisimple trace-admissible algebra then $\mathfrak{A}^{(+)}$ is semisimple and trace admissible.* It must then be true that the first step in the study of the structure of semisimple trace-admissible algebras is the case where $\mathfrak{A}^{(+)} = \mathfrak{A}$, that is, the commutative case. Let us then assume that $\mathfrak{A}$ is commutative so that we can use (6).

The mapping $a \to ax = aR_x$ is a linear transformation R_x on $\mathfrak{A}$ and it will be convenient to write yR_x^3 rather than $[(yx)x]x$. We note that for a commutative algebra $\tau(xy, z) = \tau(x, yz) = \tau(xz, y)$ and so

$$\tau(yR_x^k, z) = \tau(y, zR_x^k) \tag{9}$$

for every positive integer k. We now have $w = yx^3 - 4(yx)x^2 + (yx^2)x + 2yR_x^3 = 0$, $zw = 0$, $\tau(z, w) = \tau(zy, x^3) - 4\tau(zx^2, yx) + \tau(zx, yx^2) + 2\tau(z, yx^3) = 0$. But by (9) and the commutative law we see that $\tau(zy, x^3) + 2\tau(z, yx^3) = 4\tau(zx^2, yx) - \tau(zx, yx^2)$ is unaltered by the interchange of y and z. Thus $4\tau(zx^2, yx) - \tau(zx, yx^2) = 4\tau(yx^2, zx) - \tau(yx, zx^2)$, $5\tau(zx^2, yx) = 5\tau(zx, yx^2)$. If the characteristic of $\mathfrak{F}$ is prime to 30 we have $\tau(zx^2, yx) = \tau[z, x^2(yx)] = \tau(zx, yx^2) = \tau[z, (x^2y)x]$, $\tau[z, x^2(yx) - (x^2y)x] = 0$ for every z of $\mathfrak{A}$, $x^2(yx) - (x^2y)x$ is in the radical $\mathfrak{N} = 0$ of $\mathfrak{A}$, $x^2(yx) = (x^2y)x$. We have proved the first part of the

PRINCIPAL THEOREM. *Let $\mathfrak{A}$ be a semisimple trace-admissible power associative algebra over a field $\mathfrak{F}$ whose characteristic is prime to 30. Then $\mathfrak{A}^{(+)}$ is a semisimple Jordan algebra, $\mathfrak{A}$ is flexible, $\mathfrak{A}$ is uniquely expressible as a direct sum $\mathfrak{A} = \mathfrak{A}_1 \oplus \ldots \oplus \mathfrak{A}_t$ of simple algebras $\mathfrak{A}_i$ such that $\mathfrak{A}_i^{(+)}$ is a simple Jordan algebra. Every simple $\mathfrak{A}$ of order n over a center $\mathfrak{F}$ of characteristic 0 or $p > n$ is either a simple Jordan algebra, or has degree two, or is quasiassociative,*[8] $\mathfrak{A}^{(+)}$ is simple.

To prove the flexibility of $\mathfrak{A}$ we use the consequence

$$w = (xy + yx)x - x(xy + yx) + x^2y - yx^2 = 0 \qquad (10)$$

of the hypothesis $xx^2 = x^2x$. Then $zw = 0$ for every z of $\mathfrak{A}$, and we use (1) and (8) to obtain

$$\tau(xz - zx, xy + yx) = \tau(zy - yz, x^2).$$

Interchange z and y to obtain $\tau(xy - yx, xz + zx) = \tau(yz - zy, x^2) = \tau(xy + yx, zx - xz)$. Add $\tau(xy + yx, xz + zx)$ to both members of the resulting equation and obtain $\tau(xy, xz + zx) = \tau(xy + yx, zx)$. Then $\tau(xy, xz) + \tau(xy, zx) = \tau(xy, zx) + \tau(yx, zx)$, $\tau(xy, xz) = \tau(yx, zx)$, $\tau[x(yx) - (xy)x, z] = 0$, $(yx)x - x(xy)$ is in the radical of $\mathfrak{A}$ and $\mathfrak{A}$ is flexible.

To complete our results we note the known[9]

LEMMA 4. *Let $\mathfrak{B}$ be an ideal of a trace-admissible algebra $\mathfrak{A}$. Then the radical of $\mathfrak{B}$ is the intersection of $\mathfrak{B}$ and the radical of $\mathfrak{A}$.*

We now prove

LEMMA 5. *Let $\mathfrak{B}$ be a vector subspace of a semisimple trace-admissible algebra $\mathfrak{A}$ and $\mathfrak{B}$ be an ideal of $\mathfrak{A}^{(+)}$. Then $\mathfrak{B}$ has a unity element e, $\mathfrak{B} = \mathfrak{A}_e(1)$ is an ideal of $\mathfrak{A}$, $\mathfrak{A} = \mathfrak{B} \oplus \mathfrak{A}_e(0)$, $\mathfrak{A}^{(+)} = \mathfrak{B} \oplus \mathfrak{A}_e(0)$.*

For when $\mathfrak{A}$ is semisimple so is $\mathfrak{A}^{(+)}$. By Lemma 2 we know that $\mathfrak{B} = \mathfrak{B}_e(1)$ has a unity quantity e. Using products in $\mathfrak{A}^{(+)}$ we see that $2x \cdot e = x$ is in $\mathfrak{B}$ for every x of $\mathfrak{A}_e(1/2)$, $x \cdot e = x = 0$, $\mathfrak{B}$ is a subspace of $\mathfrak{A}_e(1)$, $x \cdot e = x$ for every x of $\mathfrak{A}_e(1)$, $\mathfrak{A}_e(1) = \mathfrak{B}$, $\mathfrak{A} = \mathfrak{A}_e(1) + \mathfrak{A}_e(0)$. Since $\mathfrak{A}$ is flexible $\mathfrak{A}_e(1)$ and $\mathfrak{A}_e(0)$ are *orthogonal* subalgebras of $\mathfrak{A}$ (and of $\mathfrak{A}^{(+)}$) and the result is proved.

The result above implies that $\mathfrak{A}$ *is simple if and only if* $\mathfrak{A}^{(+)}$ *is a simple Jordan algebra*. The remaining results of our principal theorem are trivial consequences of this result and a known[10] theorem on the structure of a flexible algebra $\mathfrak{A}$ such that $\mathfrak{A}^{(+)}$ is a simple Jordan algebra.

[1] Such algebras are the alternative algebras, right alternative algebras, Lie algebras, Jordan algebras and what the author has called standard and static algebras.

[2] A trace argument can be used in the study of alternative and associative algebras and is the only available tool for the study of Lie algebras. For its use in other cases see the following papers of the author: I. A Structure Theory for Jordan Algebras, *Ann. Math.*, **48,** 546–567 (1947); II. Power-Associative Rings, *Trans. Am. Math. Soc.*, **64,** 552–593 (1948); III. On Right Alternative Algebras, *Ann. Math.*, **50,** April number (1949).

[3] This assumption about the radical cuts out the theory of Lie algebras since all Lie algebras are nilalgebras.

[4] See II in reference 2, pages 558–561.

[5] *Ibid.*, (14) on page 556.

[6] *Ibid.*, page 564.

[7] This result is proved in II (reference 2) under an assumption which cannot be used here.

[8] Quasiassociative algebras are defined on page 583 of II (reference 2).
[9] See Theorem 2 on page 565 of II in reference 2.
[10] *Ibid.*, Theorem 13 on page 593.

THE HOMOTOPY GROUPS OF A TRIAD

BY A. L. BLAKERS AND WILLIAM S. MASSEY

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated by S. Lefschetz, April 2, 1949

In this paper we define new homotopy groups for topological spaces. These groups generalize the homotopy groups of Hurewicz. By the use of these groups and by improved methods we obtain new results about the ordinary homotopy groups, and also easier proofs of known results. Among other things, we can show that $\pi_6(S^3)$ is non-trivial.

1. One of the principal problems of modern topology is to devise methods for computing the homotopy groups of a space. The homotopy groups of even such simple spaces as spheres have not been computed except in special cases. This contrasts strongly with the situation for the homology groups, which can be computed for any triangulable space. The homotopy groups resemble homology groups in all basic properties except one: the homology groups are invariant under an excision,[1] while this is not generally true for the homotopy groups. More precisely, if a space is a union $A \cup B$ of two subspaces, then under fairly general conditions the inclusion map $i:(A, A \cap B) \to (A \cup B, B)$ (called an *excision*) induces isomorphisms of the corresponding relative homology groups, in all dimensions, but will not generally do so for the relative homotopy groups. This is perhaps the chief reason for the difficulty of computing the homotopy groups. The new homotopy groups defined in this paper are a measure of the deviation from invariance under excision for the relative homotopy groups.

We shall use the following notation for certain subsets of cartesian n-space, C^n. The coördinates of a point $x \in C^n$ are denoted by $(x_1, \ldots, x_n)$, and $|x| = (x_1^2 + \ldots + x_n^2)^{1/2}$.

$$
\begin{aligned}
E^n &= \{x \in C^n \mid |x| \leq 1\}, \\
E_-^{n-1} &= \{x \in E^n \mid |x| = 1, x_n \leq 0\}, \\
E_+^{n-1} &= \{x \in E^n \mid |x| = 1, x_n \geq 0\}, \\
S^{n-1} &= E_-^{n-1} \cup E_+^{n-1}, \qquad S^{n-2} = E_-^{n-1} \cap E_+^{n-1}, \\
E_1^n &= \{x \in E^n \mid x_2 \geq 0\}, \qquad E_2^n = \{x \in E^n \mid x_2 \leq 0\}, \\
p_0 &= (1, 0, \ldots, 0).
\end{aligned}
$$

Let $(X; A, B)$ be a *triad;* that is, A and B are subspaces of the topological space X, and $A \cap B \neq 0$. Choose a base point $x_0 \in A \cap B$. A *map*

$$f:(E^n; E^{n-1}_+, E^{n-1}_-, p_0) \to (X; A, B, x_0)$$

is a continuous function $f:E^n \to X$ such that $f(E^n_+) \subset A$, $f(E^n_-) \subset B$, and $f(p_0) = x_0$. Two such maps f_0, f_1, are *homotopic* if they are connected by a continuous 1-parameter family of maps,

$$f_t:(E^n; E^{n-1}_+, E^{n-1}_-, p_0) \to (X; A, B, x_0),$$

where $0 \leq t \leq 1$.

LEMMA. *If $n \geq 3$, and $f:(E^n; E^{n-1}_+, E^{n-1}_-, p_0) \to (X; A, B, x_0)$, then there exist homotopic maps f', f'', with $f'(E^n_1) = f''(E^n_2) = x_0$.*

Denote the set of all homotopy classes of mappings $(E^n; E^{n-1}_+, E^{n-1}_-, p_0) \to (X; A, B, x_0)$ by $\pi_n(X; A, B, x_0)$. For $n \geq 3$, we define an addition between any two elements $\alpha, \beta \in \pi_n(X; A, B, x_0)$ as follows: Choose maps f, g belonging to the homotopy classes α, β, respectively, with $f(E^n_2) = g(E^n_1) = x_0$. Define $h:(E^n; E^{n-1}_+, E^{n-1}_-, p_0) \to (X; A, B, x_0)$ by $h|E^n_1 = f|E^n_1$ and $h|E^n_2 = g|E^n_2$. Then $\alpha + \beta$ is defined to be the homotopy class of h. With this definition $\pi_n(X; A, B, x_0)$ becomes a group, called the *n*th *homotopy group of the triad* $(X; A, B)$ *at the base point* x_0.

Just as in the case of the relative homotopy groups, a continuous map $f:(X; A, B, x_0) \to (X'; A', B', x_0')$ induces homomorphisms of the triad homotopy groups. We denote these by $f_*:\pi_n(X; A, B, x_0) \to \pi_n(X'; A', B', x_0')$.

Associated with the triad $(X; A, B)$ are two *boundary homomorphisms,*

$$\partial:\pi_n(X; A, B, x_0) \to \pi_{n-1}(A, A \cap B, x_0),$$
$$\partial':\pi_n(X; A, B, x_0) \to \pi_{n-1}(B, A \cap B, x_0),$$

defined as follows. If $f:(E^n; E^{n-1}_+, E^{n-1}_-, p_0) \to (X; A, B, x_0)$ represents $\alpha \in \pi_n(X; A, B, x_0)$, $(n \geq 3)$, then $f|E^{n-1}_+$ represents $\partial(\alpha)$, and $f|E^{n-1}_-$ represents $\partial'(\alpha)$. Now consider the following sequence of groups and homomorphisms:

$$\ldots \xrightarrow{\partial} \pi_n(A, A \cap B) \xrightarrow{i_*} \pi_n(X, B) \xrightarrow{j_*} \pi_n(X; A, B) \xrightarrow{\partial} \pi_{n-1}(A, A \cap B) \xrightarrow{i_*} \ldots$$

The homomorphisms i_* and j_* are induced by the inclusion maps

$$i:(A, A \cap B) \to (X, B), \qquad j:(X; x_0, B) \to (X; A, B).$$

This sequence is one of the two homotopy sequences of the triad $(X; A, B)$; the other sequence is obtained by interchanging the roles of A and B throughout.

THEOREM 1. *The homotopy sequences of a triad are both exact.*

If $X = A \cup B$, then the inclusion map i is an excision. It follows from

exactness that i_* is an isomorphism in all dimensions if and only if $\pi_n(A \cup B; A, B) = 0$ for all n. Thus if $\pi_n(A \cup B; A, B) \neq 0$ for some integer n, then the excision i cannot induce isomorphisms onto in all dimensions, and the groups $\pi_n(A \cup B; A, B)$ are a "measure" of the amount of deviation from invariance under excision.

As an example of the application of these results, consider the triad $(S^n; E_+^n, E_-^n)$. It can be shown (see below) that $\pi_p(S^n; E_+^n, E_-^n) = 0$ for $p < 2n - 1$, and that $\pi_{2n-1}(S^n; E_+^n, E_-^n)$ is infinite cyclic. In this case the excision homomorphism $i_*: \pi_p(E_+^n, S^{n-1}) \to \pi_p(S^n, E_-^n)$ is equivalent to Freudenthal's "Einhängung";[2] more precisely, in the following diagram,

$$\begin{array}{ccc} & & \pi_{p-1}(S^{n-1}) \xrightarrow{E} \pi_p(S^n) \\ & & \partial \uparrow \qquad\quad \downarrow k_* \\ \ldots \to \pi_p(E_+^n, S^{n-1}) & \xrightarrow{i_*} & \pi_p(S^n, E_-^n) \to \ldots \end{array}$$

the boundary homomorphism ∂, and the homomorphism k_* induced by an inclusion map, are isomorphisms onto, and the commutativity relation $i_* = k_* E \partial$ holds, where E is the Einhängung. These facts, together with the exactness of the homotopy sequence of the triad, imply the Freudenthal theorems.

2. The homotopy groups of triads have many properties which are the analogs of familiar properties of the homotopy groups of spaces or pairs of spaces. We now list some of these properties:

(*a*) $\pi_n(X; A, B, x_0)$ is abelian for $n > 3$; simple examples show that it need not be so for $n = 3$.

(*b*) The system of groups $\pi_n(X; A, B, x)$ for $x \in A \cap B$ forms a local system of groups[3] in the space $A \cap B$ in the usual sense, and $\pi_1(A \cap B, x_0)$ is a group of operators on $\pi_n(X; A, B, x_0)$.

(*c*) If $A \supset B$, then $\pi_n(X; A, B)$ is isomorphic to $\pi_n(X, A)$, and the homotopy sequence of the triad $(X; A, B)$ reduces to that of the triple (X, A, B).

(*d*) In case $\pi_3(X; A, B)$ is non-abelian, the kernels of the boundary homomorphisms $\partial: \pi_3(X; A, B) \to \pi_2(A, A \cap B)$ and $\partial': \pi_3(X; A, B) \to \pi_2(B, A \cap B)$ are both contained in the center of $\pi_3(X; A, B)$.

3. A topological space X is said to be *n-connected* ($n > 0$) if it is arcwise connected and $\pi_p(X) = 0$ for $1 \leq p \leq n$. A pair (X, A) is said to be *n-connected* ($n \geq 1$) if both X and A are arcwise connected, the natural homomorphism $\pi_1(A) \to \pi_1(X)$ is a homomorphism onto, and $\pi_p(X, A) = 0$, $2 \leq p \leq n$. Similarly, we will say a triad $(X; A, B)$ is *n-connected* ($n \geq 2$) if both of the pairs $(A, A \cap B)$ and $(B, A \cap B)$ are 1-connected, and $\pi_p(X; A, B) = 0$, $2 \leq p \leq n$. By $\pi_2(X; A, B) = 0$ we mean that this set of homotopy classes of mappings contains a single element, the homotopy class of the constant map into x_0.

Let (X^*, X) be a pair, and $\psi:(E^n, S^{n-1}) \to (X^*, X)$, $(n > 1)$, a continuous map which is a homeomorphism of $(E^n - S^{n-1})$ onto $X^* - X$. Let $\mathcal{E}^n = \psi(E^n)$, $\dot{\mathcal{E}}^n = \mathcal{E}^n \cap X = \psi(S^{n-1})$. We say that the space X^* is obtained from X by *adjunction of the cell* $\mathcal{E}^n$. Note that $(X^*, X, \mathcal{E}^n)$ is a triad, and $\dot{\mathcal{E}}^n$ is 0-connected.

THEOREM 2. *If the pair $(X, \dot{\mathcal{E}}^n)$ is m-connected, $m \geq 1$, then the triad $(X^*, X, \mathcal{E}^n)$ is $(m + n - 1)$-connected. (In case $m = 1$, it is also necessary to assume that $\pi_2(X, \dot{\mathcal{E}}^n)$ is abelian; in case $n = 2$, it is necessary to assume that X is simple relative to $\dot{\mathcal{E}}^n$ in all dimensions.)*

THEOREM 3. *If the space $\dot{\mathcal{E}}^n$ is m-connected, $m \geq 1$, then the boundary homomorphism*

$$\partial:\pi_{i+1}(X^*, X, \mathcal{E}^n) \to \pi_i(\mathcal{E}^n, \dot{\mathcal{E}}^n)$$

is the zero homomorphism for $2 \leq i \leq m + n - 1$.

The proofs of theorems 2 and 3 are a straightforward application of the theory of "obstructions"[4] to extensions and deformations of continuous mappings. By combining these theorems with the exactness of the homotopy sequence of the triad $(X^*, X, \mathcal{E}^n)$, important results are obtained as corollaries. Among these are the previously mentioned assertion that $\pi_p(S^n; E^n_+, E^n_-) = 0$ for $p < 2n - 1$, and an essential generalization of a theorem of J. H. C. Whitehead[5] about the homotopy groups of the pair (X^*, X).

4. Let $f:(E^n; E^{n-1}_+, E^{n-1}_-) \to (X; A, B)$ be a map, and assume $u \in H^p(X, A)$, $v \in H^q(X, B)$, $p + q = n + 1$ and $u \smile v = 0$, where $u \smile v$ is the Alexander-Cech-Whitney cup product. In a recent paper,[6] Steenrod has shown how to associate with the elements u, v an element $u \smile_f v$ of the factor group $H^n(E^n, S^{n-1})/f^*H^n(X, A \cup B)$. He has proved that the pairing thus defined, called the *functional cup product* is bilinear, and depends only on the homotopy class of f. Hence if $\alpha \in \pi_n(X; A, B, x_0)$ is the homotopy class of f, we can write $u \smile_f v = u \smile_\alpha v$ without ambiguity.

THEOREM 4. *If $H^n(X, A \cup B) = H^{n+1}(X, A \cup B) = 0$, and $\gamma = \alpha + \beta$, $\alpha, \beta \in \pi_n(X; A, B, x_0)$, $n > 2$, then*

$$u \smile_\alpha v + u \smile_\beta v = u \smile_\gamma v.$$

This conclusion can be alternately stated: *The functional cup product determines a homomorphism of the group $\pi_n(X; A, B)$ into the group of bilinear functions*

$$[H^p(X, A), H^q(X, B)] \to H^n(E^n, S^{n-1}).$$

In the important special case where $(X; A, B) = (S^n; E^n_+, E^n_-)$, and we use cohomology groups with integer coefficients, additional results can be obtained. It is easily seen that the group of bilinear maps

$$[H^n(S^n, E_+^n), H^n(S^n, E_-^n)] \to H^{2n-1}(E^{2n-1}, S^{2n-2})$$

is an infinite cyclic group. Denote this group by G_n. Then we can prove

THEOREM 5. *The homomorphism* $\pi_{2n-1}(S^n; E_+^n, E_-^n) \to G_n$ *determined by the functional cup product is an isomorphism onto.*

Now consider the following portion of the homotopy sequence of the triad $(S^n; E_+^n, E_-^n)$:

$$\pi_{2n-1}(S^n, E_-^n) \xrightarrow{j_*} \pi_{2n-1}(S^n; E_+^n, E_-^n) \xrightarrow{\partial} \pi_{2n-2}(E_+^n, S^{n-1}) \xrightarrow{i_*} \pi_{2n-2}(S^n, E_-^n).$$

Using the known properties of the functional cup product, additional information can be obtained about the homomorphisms j_*, ∂, and i_*.

(*a*) If n is odd, then $j_* = 0$, ∂ is an isomorphism into, and image ∂ is an infinite cyclic subgroup of $\pi_{2n-2}(E_+^n, S^{n-1})$.

(*b*) If n is even, and $\pi_{2n-1}(S^n)$ has an element of Hopf Invariant 1, then j_* is onto, $\partial = 0$, and i_* is an isomorphism onto. This case is known to occur for $n = 2, 4, 8$.

(*c*) If n is even and $\pi_{2n-1}(S^n)$ has no element of Hopf Invariant 1, then it must have an element of Hopf Invariant 2. In this case j_* is a homomorphism onto the subgroup of $\pi_{2n-1}(S^n; E_+^n, E_-^n)$ of index two, and image ∂ is a cyclic subgroup of $\pi_{2n-2}(E_+^n, S^{n-1})$, of order two. G. W. Whitehead[7] has recently announced that this case occurs when $n \equiv 2 \bmod 4$, $n > 2$.

These three statements are essentially the Freudenthal theorems, together with some improvements due to G. W. Whitehead.[7]

5. Let A and B be arcwise connected spaces, $a_0 \in A$, $b_0 \in B$, and let $A \vee B$ be the subspace of the cartesian product $A \times B$ defined by $A \vee B = (A \times b_0) \cup (a_0 \times B)$. $A \vee B$ is the union of A and B with a single point in common. Define $\mu_1 : A \to A \vee B$ by $\mu_1(x) = (x, b_0)$ and $\mu_2 : B \to A \vee B$ by $\mu_2(y) = (a_0, y)$, for $x \in A$, $y \in B$. Consider the following homomorphisms $(n > 1)$:

$$\mu_{1*} : \pi_n(A) \to \pi_n(A \vee B),$$
$$\mu_{2*} : \pi_n(B) \to \pi_n(A \vee B),$$
$$\partial : \pi_{n+1}(A \times B, A \vee B) \to \pi_n(A \vee B).$$

We can paraphrase a result of G. W. Whitehead[7] as follows: The homomorphisms μ_{1*}, μ_{2*}, and ∂ are isomorphisms into, and $\pi_n(A \vee B)$ splits up into the direct sum

$$\mu_{1*}\pi_n(A) + \mu_{2*}\pi_n(B) + \partial\pi_{n+1}(A \times B, A \vee B).$$

From this it follows by an easy argument using exact sequences, that for $n > 2$, $\pi_n(A \vee B; A, B)$ is isomorphic to $\pi_{n+1}(A \times B, A \vee B)$, and that in the homotopy sequence of the triad $(A \vee B; A, B)$, the boundary operators are both trivial in all dimensions. If A and B are spheres, further

results can be obtained. If theorems 2 and 3 are applied with $X^* = S^p \times S^q$, and $\mathcal{E}^n$ a $(p+q)$-simplex in a triangulation of $S^p \times S^q$, we find that $\pi_i(S^p \times S^q, S^p \vee S^q) \approx \pi_i(E^{p+q}, S^{p+q-1})$ for $i \leq p+q+\min(p,q)-2$. It follows that $\pi_j(S^p \vee S^q; S^p, S^q) \approx \pi_j(S^{p+q-1})$ for $j \leq p+q+\min(p,q)-3$.

Let $\varphi_n : (S^n; E^n_+, E^n_-) \to (S^n_1 \vee S^n_2; S^n_1, S^n_2)$ be the map defined by identifying all points of the "equator" $S^{n-1} \subset S^n$ to a single point. Then φ_n induces a homomorphism of the corresponding triad homotopy sequences, as shown in the following diagram:

$$\begin{array}{ccccccccc} \cdots \xrightarrow{\partial} \pi_r(E^n_+, S^{n-1}) & \xrightarrow{i_*} & \pi_r(S^n, E^n_-) & \xrightarrow{j_*} & \pi_r(S^n; E^n_+, E^n_-) & \xrightarrow{\partial} & \cdots \\ \downarrow & & \downarrow & & \downarrow \varphi_{n_*} & & \\ \cdots \xrightarrow{\partial} \pi_r(S^n_1) & \to & \pi_r(S^n_1 \vee S^n_2, S^n_2) & \to & \pi_r(S^n_1 \vee S^n_2; S^n_1, S^n_2) & \xrightarrow{\partial} & \cdots \end{array}$$

Now $\pi_r(S^n) \approx \pi_r(S^n, E^n_-)$, and if $r \leq 3n-3$, $\pi_r(S^{2n-1}) \approx \pi_r(S^n_1 \vee S^n_2; S^n_1, S^n_2)$ by the preceding paragraph. The composition of these two isomorphisms with the homomorphism $\varphi_{n_*} j_*$ of the diagram yields a homomorphism

$$H : \pi_r(S^n) \to \pi_r(S^{2n-1}), \ (r \leq 3n-3).$$

It may be shown that this homomorphism is the same as the Generalized Hopf Homomorphism defined by G. W. Whitehead.[7] However, this definition is simpler and easier to work with than the original definition. Furthermore, we have demonstrated the existence of H for $r \leq 3n-3$, instead of for $r \leq 3n-4$, as was originally done by Whitehead. By using Whitehead's methods, we can show that in the limiting case for $n = 3$, the homomorphism

$$H : \pi_6(S^3) \to \pi_6(S^5)$$

is a homomorphism onto. Since $\pi_6(S^5)$ is known to be cyclic of order two, this shows that $\pi_6(S^3) \neq 0$. An essential map of S^6 onto S^3 may be constructed as follows: Let $f : S^3 \times S^2 \to S^2$ be a map of type[8] (α, ι), where α is a generator of $\pi_3(S^2)$ and $\iota \in \pi_2(S^2)$ is the class of the identity map. Now apply the Hopf construction[9] to f to obtain a map $F : S^6 \to S^3$.

[1] Eilenberg, S., and Steenrod, N. E., these PROCEEDINGS, **31**, 117–120 (1945). Axiom 6 of this note is called the "excision axiom."

[2] Freudenthal, H., *Compositio Math.*, **5**, 299–314 (1937).

[3] Steenrod, N. E., *Ann. Math.*, **44**, 610–627 (1943).

[4] See Eilenberg, S., *Ann. Math.*, **41**, 231–251 (1940), and Blakers, A. L., *Bull. Am. Math. Soc.* **54**, abstract no. 413 (1948). Also a forthcoming paper by Olum, P., in *Ann. Math.* entitled "Obstructions to Extensions and Homotopies."

[5] Whitehead, J. H. C., *Proc. Lond. Math. Soc.*, **48**, 265, theorem 8 (1944).

[6] Steenrod, N. E., these PROCEEDINGS, **33**, 124–128 (1947). A complete exposition will appear soon in the *Ann. Math.*

[7] Whitehead, G. W., these PROCEEDINGS, **32,** 188–190 (1946); also *Bull. Am. Math. Soc.*, **54,** abstract no. 546 (1948). A complete exposition of these results will appear in *Ann. Math.* The authors are indebted to Professor Whitehead for allowing them access to his manuscript before publication.

[8] Whitehead, G. W., *Ann. Math.*, **47,** 460–475 (1946).

[9] Hopf, H., *Fund. Math.*, **25,** 427–440 (1935); also, Whitehead, G. W., *Ann. Math.*, **43,** 636 (1942).

---

*ON THE "CENTRAL" PROBABILITY PROBLEM**

BY MICHEL LOÈVE

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT BERKELEY

Communicated by G. C. Evans, April 4, 1949

Let $\mathcal{L}(X)$, $\mathcal{L}(Y)$ be the probability laws of the real random variables (r. v.) X, Y, laws characterized either by the distribution functions (d. f.) $F(x)$, $G(x)$ or by their Fourier-Stieljes transforms—characteristic functions (c. f.) $f(u)$, $g(u)$. When X or Y possess subscripts, their d. f. and c. f. will have the same subscripts.

I. The Problem.—The problem of convergence towards a normal law of $\mathcal{L}\left(\sum_{k=1}^{n} X_k/c_n\right)$ when $n \to \infty$, X_k's being independent r. v. and $c_n \to \infty$ being "sur" numbers, has played a central rôle in probability theory. It has been finally solved in 1934–1935 by Feller[1] and by P. Lévy.[8]

In its present formulation the "central" problem is that of limit laws, when $n \to \infty$, of sums

($\mathcal{P}$) $X_n = X_{n,1} + X_{n,2} + \ldots + X_{n,\nu_n}$, $\nu_n \to \infty$, under the assumption of *independence*, i.e.

($\mathcal{I}$) for every fixed n, $X_{n,1}$, $X_{n,2}$, $\ldots$, X_{n,ν_n} are independent, and of *asymptotically and uniformly negligible components:*

($\mathfrak{N}$) $\max_{1 \leq k \leq \nu_n} P(|X_{n,k}| > \epsilon) \to 0$ for every given $\epsilon > 0$.

It has been shown that the class of limit laws coincides with that of infinitely divisible (i. d.) laws (Khintchine,[7] also Lévy,[8] after partial results of Bawly[1]) characterized, roughly speaking, as convolutions of a normal law and a continuum of Poisson laws (P. Lévy[8] after partial results of Kolmogoroff).[6] Conditions for convergence towards any i. d. law were given by Gnedenko[5] and Doeblin.[3]

The assumption ($\mathfrak{N}$) is natural if one wants the limit laws to depend primarily upon the fact that the number ν_n of components of sums X_n would be infinitely increasing. But the assumption ($\mathcal{I}$) is very restrictive. When ($\mathcal{I}$) is not assumed, the fundamental result for the primitive form of the problem is due to S. Bernstein:[2] if $E|X_k|^3 < \infty$, normal con-

vergence holds under conditions which, in the case of independence, reduce to those of Liapounoff; for the general limiting process ($\mathcal{V}$), without any assumption about the moments, conditions which under ($\mathcal{I}$) reduce to those of Feller were given by the author.[9a]

The general problem of limit laws of sums in the case of dependence has not been considered until now and this work is an attempt in this direction. It appeared useful to extend the whole "central" problem to that of "asymptotic equivalence" (as. eq.) of laws, to be defined. This, in turn, led to the introduction of r. v. "in a wide sense," i.e., with possible positive probabilities for infinite values. In general, laws of two sums, varying with n, are compared for $n \gg 1$; if one of them reduces to a fixed law or converges towards it, the problem reduces to that of limit laws. In the case of one variable law $\mathcal{L}(X_n)$ it becomes that of the asymptotic behavior of $\mathcal{L}(X_n)$: its asymptotic type, if any, and more generally types of its accumulation laws, classes of as. eq. laws, etc.... The tools used, and some applications directly related to the problem of limit laws, which contain results obtained under ($\mathcal{I}$), are stated here.

It should be noticed that S. Bernstein's and the author's previous conditions for normal convergence (see also[9b]) appear as unnecessarily restrictive.

II. The Tools.—Let H-functions: $H(x)$, $H_n(x)$, ... be of uniformly bounded variation over $(-\infty, +\infty)$ and denote their Fourier-Stieljes transforms by $h(u)$, $h_n(u)$, ... Because of the meaning to be attached only to $\Delta H(x)$ and not to $H(x)$ itself, and the fundamental rôle of Fourier-Stieljes transforms, H-functions which, up to a constant, coincide at all their continuity points will be considered as equivalent and represented by the normalized elements of their equivalence classes:

$$H(-\infty) = \lim_{x \to -\infty} H(x) = 0, \qquad H(x) = \frac{H(x-0) + H(x+0)}{2}.$$

Any convergence of H-functions will be defined up to an equivalence. If not otherwise stated, limits will be considered for $n \to \infty$ and for all (finite) values of the arguments.

When $H_n(x)$ converges towards a limit function, this is an H-function: write $H_n(x) \to H(x)$—"simple" convergence. If, moreover, $\lim_{n \to \infty} H_n(+\infty) = H(+\infty)$, i.e., $\lim_{n \to \infty} \lim_{x \to +\infty} H_n(x) = \lim_{x \to +\infty} \lim_{n \to \infty} H_n(x)$, write $H_n(x) \xrightarrow{c} H(x)$—"complete" convergence (not to be confused with the one where var. $H_n \to$ var. H). The tools used are the following four lemmas of which A and B are extensions of the continuity theorem for c. f., C is immediate and the basic lemma D is elementary.

LEMMA A. *If $h_n(u)$ converges towards $k(u)$ almost everywhere then $H_n(x) \to$ some $H(x)$ and $h_n(u) = k(u)$ almost everywhere.*

LEMMA B. *If $h_n(u)$ converges towards $k(u)$ continuous at the origin then $H_n(x) \xrightarrow{c} H(x)$.*

Let S be a bounded set on the real line and S' its complement with respect to $(-\infty, +\infty)$.

LEMMA C.

$$|h(u)| \leqq |H(+\infty)| + 2\int_{S'} |dH(x)| + \sum_{j=1}^{m} \frac{|u|^j}{j!} \left| \int_S x^j dH(x) \right| + \frac{|u|^{m+1}}{j^{m+1}} \int_S |x^{m+1} dH(x)|.$$

Let $X_n = X_{n,1} + \ldots + X_{n,\nu_n}$, $Y_n = Y_{n,1} + \ldots + Y_{n,\nu_n}$, $Z_{n,k} = X_{n,1} + \ldots + X_{n,k-1} + Y_{n,k+1} + \ldots + Y_{n,\nu_n}$ and the prime denote conditional expressions given $Z_{n,k}$.

LEMMA D. $|f_n(u) - g_n(u)| \leqq \sum_{k=1}^{\nu_n} E|f'_{n,k}(u) - g'_{n,k}(u)|$.

III. Asymptotic Equivalence.—If $H_n(x) = F_n(x) - G_n(x) \xrightarrow{c} 0$ we say that $\mathfrak{L}(X_n)$ and $\mathfrak{L}(Y_n)$ are (completely) as. eq. and write $\mathfrak{L}(X_n) \overset{c}{\sim} \mathfrak{L}(Y_n)$. Put

$$a'_{n,k}(\epsilon) = \int_{|x|<\epsilon} x dF'_{n,k}(x), \quad \sigma'^2_{n,k}(\epsilon) = \int_{|x|<\epsilon} [x - a'_{n,k}(\epsilon)]^2 dF'_{n,k}(x)$$

and $b'_{n,k}(\epsilon)$, $\tau'^2_{n,k}(\epsilon)$ when $F'_{n,k}(x)$ is replaced by $G'_{n,k}(x)$. Suppress the primes when they disappear in $F'_{n,k}(x)$, respectively, in $G'_{n,k}(x)$.

We state two results directly useful for the problem of limit laws and, to simplify, consider only usual—"strict" r. v.

I. *If, for $\epsilon \to 0$*

$$\overline{\lim_{n\to\infty}} \sum_k E \int_{|x|\geq\epsilon} |dF'_{n,k}(x) - dG'_{n,k}(x)| = o(1),$$

$$\overline{\lim_{n\to\infty}} \sum_k E|a'_{n,k}(\epsilon) - b'_{n,k}(\epsilon)| = o(1), \quad \overline{\lim_{n\to\infty}} \sum_k E|\sigma'^2_{n,k}(\epsilon) - \tau'^2_{n,k}(\epsilon)| = o(1), \quad \overline{\lim_{n\to\infty}} \sum_k E\tau'^2_{n,k}(\epsilon) = o(\epsilon^{-1})$$

then $\mathfrak{L}(X_n) \overset{c}{\sim} \mathfrak{L}(Y_n)$.

Take $\mathfrak{L}'(Y_{n,k})$ to be the "conditional i. d. law" defined by

$$\log g'_{n,k}(u) = ia'_{n,k}(\epsilon)u + \int_{-\infty}^{+\infty} (\exp.\{iu[x - a'_{n,k}(\epsilon)]\} - 1) dF'_{n,k}(x)$$

II. *If, for $\epsilon \to 0$,*

$$\overline{\lim_{n\to\infty}} \sum_k E\left| \int_{-\infty}^{+\infty} (\exp.\{iu[x - a'_{n,k}(\epsilon)]\} - 1) dF'_{n,k}(x) \right|^2 = o(1)$$

then

$$\overline{\lim_{n \to \infty}} \left| f_n(u) - g_n(u) \right| = o(1).$$

Under $(\mathfrak{K})$, *if*

$$\sum_k \int_{|x| \geqq \epsilon} dF_{n,k}(x) \leqq C_1 < \infty, \qquad \sum_k E\sigma'^2_{n,k}(\epsilon) \leqq C_2 < \infty$$

then $\mathfrak{L}(X_n) \underset{c}{\sim} \mathfrak{L}(Y_n)$.

Applied to the problem of limit laws these results give:

Asymptotic independence. If the hypotheses of I are satisfied with $G'_{n,k}(x) = F_{n,k}(x)$ *then, under* $(\mathfrak{K})$, *the necessary and sufficient conditions for* $\mathfrak{L}(X_n)$ *to converge towards a law, necessarily i. d. are, that for any* $\epsilon > 0$ $\sum_k a_{n,k}(\epsilon) \to a(\epsilon)$

$\lim_{\epsilon \to 0} \sum_k \sigma^2_{n,k}(\epsilon) = \sigma^2$ *and*

$$\sum_k \int_{-\infty}^{-\epsilon} dF_{n,k}(x) \to \int_{-\infty}^{-\epsilon} dL(x), \quad \sum_k \int_{\epsilon}^{\infty} dF_{n,k}(x) \to \int_{\epsilon}^{\infty} dL(x)$$

where $a(\epsilon)$, σ^2 *and* $L(x)$ *characterize the i. d. limit law.* In the case of independence this becomes the Gnedenko-Doeblin result.

Let us give two simple consequences; $\epsilon > 0$ is arbitrarily small.

Examples: 1. If

$$\sum_k \int_{|x| \geqq \epsilon} dF_{n,k}(x) \to 0,$$

$$\sum_k E\left| a'_{n,k}(\epsilon) - a_{n,k}(\epsilon) \right| \to 0, \sum_k a_{n,k}(\epsilon) \to a$$

$$\sum_k E\left| \sigma'^2_{n,k}(\epsilon) - \sigma^2_{n,k}(\epsilon) \right| \to 0, \qquad \sum_k \sigma^2_{n,k}(\epsilon) \to \sigma^2$$

then $\mathfrak{L}(X_n)$ *converges towards the normal law* $N(a, \sigma^2)$—contains the S. Bernstein's and author's conditions.

Let $S = (-\epsilon, +\epsilon) \cup (1 - \epsilon, 1 + \epsilon)$.

2. If $\sum_k E \int_{|x-1|<\epsilon} \left| dF'_{n,k}(x) - dF_{n,k}(x) \right| \to 0, \sum_k \int_{|x-1|<\epsilon} dF_{n,k}(x) \to \lambda,$ $\sum_k E\left| a'_{n,k}(\epsilon) \right| \to 0, \sum_k E\sigma'^2_{n,k}(\epsilon) \to 0, \sum_k \int_{S'} dF_{n,k}(x) \to 0,$ *then* $\mathfrak{L}(X_n)$ *converges towards the Poisson law with parameter* λ—in the case of independence these conditions reduce to those of Gnedenko-Marcinkiewicz.

IV. Extensions.—1. Let now ν_n be a r. v. Put $P(\nu_n = r) = p_n(r)$, $P(\nu_n \leqq r) = P_n(r)$, denote by the subscript r the expectations given $\nu_n = r$, assume that the quantities below exist (they certainly do when $E\nu_n < \infty$), and write $\mathcal{E}_{n,k}$ for $\sum_{r \geqq k} p_n(r) E_r$. Notice that to the condition $\nu_n \to \infty$ when ν_n is "sur" would correspond, for instance, the condition: $\mathfrak{L}(Y_n)$ converges towards a law (in [illegible] sense) degenerated at infinity.

The results above hold for ν_n random if we replace E by $\mathcal{E}_{n,k}$ (observe that $F_{n,k}(x) = EF'_{n,k}(x)$).

If the $X_{n,k}$'s do not depend upon ν_n, then we have only to replace them by the r. v. in a wide sense with conditional laws $P_n(k)F'_{n,k}(x)$ instead of $F'_{n,k}(x)$.

2. The results above hold for random vectors $X_{n,k}$, $Y_{n,k}$' . . . with the corresponding interpretation of the symbols used.

It is hoped that the proofs and various details will be published in the Statistical Series of the University of California.

* This work has been supported in part by the Office of Naval Research.

[1] Bawly, G., *Matem. Sbornik.*, **1** (43), 917–929 (1936).

[2] Bernstein, S., *Math. Ann.*, **97,** 1–59 (1927).

[3] Doeblin, W., *Bull. Sc. Math.*, **63,** 22–32, 35–64 (1939).

[4] Feller, W., *Math. Zeitschr.*, **40,** 521–539 (1935); **42,** 301–312 (1937).

[5] Gnedenko, B., *DOKL.* **18,** 231–234 (1938); **22,** 886–871 (1939). *IZV. A. N.* 181–232, 643–647 (1939).

[6] Kolmogoroff, A., *Rend. Lincei,* **15,** 805–808, 866–869 (1932).

[7] Khintchine, A., *Matem. Sbornik,* **2** (44), 79–119 (1937).

[8] Lévy, P., *Théorie de l'Addition des Variables Aléatoires,* Paris, Gauthier-Villars (1937).

[9] Loève, M., *J. de Math.,* **24,** 249–318 (1945).

[9'] Loève, M., *C. R. Ac. Sc. Paris,* **227,** 1335–1337 (1948).

[10] Marcinkiewicz, J., *Fund. Math.*, **30,** 354 (1938).

ON CONGRUENCES WHICH RELATE THE FERMAT AND WILSON QUOTIENTS TO THE BERNOULLI NUMBERS

By H. S. Vandiver

Department of Applied Mathematics, University of Texas

Communicated April 30, 1949

In a recent paper[1] the writer proved a result which may be written in the following form

$$-q\left(\frac{m}{r}\right) + W - \sum_{n=1}^{p-2}\left(\frac{m}{r}\right)^n \frac{b_n}{n} \equiv \frac{\lambda(m, r)}{r}. \qquad (1)$$

modulo p, where $m > 0$, $0 < r < p$, $(m, p) = 1$, $(k, p) = 1$, p is an odd prime,

$$q(s) = \frac{s^{p-1} - 1}{p},$$

$$W = \frac{(p-1)! + 1}{p},$$

$$\lambda(m, r) = \sum{}' \frac{\rho^r - 1}{\rho^p - 1};\ m > 1,$$

with $\lambda(1, r) = 0$, ρ an mth root of unity $\neq 1$, the sign $\sum'$ means summation over all mth roots of unity $\neq 1$, and b_n is the Bernoulli number defined by the recursion formula $(b + 1)^k = b_k$ for every $k > 1$, where the left-hand member is expanded and b_i substituted for b^i in the result.

In the present paper we deduce some new relations concerning the Fermat and Wilson quotients and the Bernoulli numbers. The principal results are (6a), (9), (11), (11a) and (22).

To show that (1) is equivalent to the other relation referred to, we note that the $\beta(m, r)$ of our paper *C* satisfies

$$\lambda(m, r) = \beta(m, r) - \sum' \frac{1}{\rho^p - 1} \tag{2}$$

and applying the known relation

$$\sum' \frac{1}{\rho^p - 1} = \sum' \frac{1}{\rho - 1} = \sum' \frac{\rho + 2\rho^2 + \ldots + (m-1)\rho^{m-1}}{m},$$

gives

$$\sum' \frac{1}{\rho^p - 1} = \frac{1 - m}{2} = \frac{1}{2} - \frac{m}{2},$$
$$= \frac{1}{2} - (-mb_1).$$

We also have

$$q\left(\frac{m}{r}\right) = q(m) - q(r),$$

by the definition of q. We now consider some of the properties of $\lambda(m, r)$. Obviously

$$\lambda(m, r) = \lambda(m, r + tm). \tag{3}$$

Set $ps \equiv r \pmod m$, $m > 1$, $0 < s < m$, then

$$\lambda(m, r) = \sum' \frac{\rho^r - 1}{\rho^p - 1} = \sum' \frac{\rho^{ps} - 1}{\rho^p - 1},$$
$$= \sum' \frac{\rho^s - 1}{\rho - 1},$$
$$= m - s; \quad m > 1. \tag{4}$$

Also as r ranges over the integers $1, 2, \ldots, m - 1$, then the corresponding numbers s range over the same set in some order. Hence

$$\sum_{r=1}^{m-1} \lambda(m, r) = (m-1)\left(m - \frac{m(m-1)}{2}\right),$$

$$= \frac{m(m-1)}{2}. \tag{5}$$

In particular if $p \equiv 1 \pmod{m}$, $\lambda(m, r) = m - r$. Also $\lambda(m, r) = 0$ if $r \equiv 0 \pmod{m}$.

We now deduce some relations from (1). Set $m = 1$ and $r = 1, 2, \ldots, p-1$, in turn and add the resulting congruences, then since

$$\sum_{m=1}^{p-1} r^a \equiv 0 \pmod{p},\ a < p-1, \tag{5a}$$

we have the known relation

$$W \equiv \sum_{r=1}^{p-1} q(r) \pmod{p}. \tag{6}$$

Now set in (1), $r = 1$ and add the relations found by setting $m = 1, 2, \ldots, p-1$ in turn we have[2] using (5a) and (6)

$$-2W \equiv \sum_{m=1}^{p-1} \lambda(m, 1) \pmod{p}. \tag{6a}$$

Multiplying (1) through by r, the result may be written

$$-rq(m) + rq(r) + rW - \sum m^n r^{p-n} \frac{b_n}{n} \equiv \lambda(m, r) \pmod{p}. \tag{7}$$

Multiply this through by m^k, and add the results for $m = 1, 2, \ldots, p-1$ we find from (5a), using the known relation (C, p. 109),

$$\sum_{a=1}^{p-1} a^t q(a) \equiv -\frac{b_t}{t} \pmod{p}, \tag{8}$$

the congruence

$$-r\frac{b_k}{k} - r^{k+1}\frac{b_{p-1-k}}{p-1-k} \equiv \sum_{m=1}^{p-1} m^k \lambda(m, r) \pmod{p}. \tag{9}$$

Now, taking $m = 1, 2, \ldots, p-1$ in (7) as it stands, and adding, we have by (6) a result which may be written

$$q(r) + 2W \equiv -\sum_{m=1}^{p-1} \frac{\lambda(m, r)}{r} \pmod{p}, \tag{10}$$

and by using (6a) we have

$$q(r) \equiv -\sum_{m=1}^{p-1} \frac{\lambda(m, r)}{r} + \sum_{m=1}^{p-1} \lambda(m, 1) \pmod{p}. \tag{11}$$

On the surface, at least, this relation throws little if any light on the properties of $q(r)$, but on the other hand the known properties of $q(r)$ enable us to find relations involving $\lambda(m, r)$ alone for various values of m and r. For example since, if $rt < p$,

$$q(rt) \equiv q(r) + q(t) (\text{mod } p),$$

we have from (11),

$$\sum_{m=1}^{p-1} \left(\frac{\lambda(m, r)}{r} + \frac{\lambda(m, t)}{t} \right) \equiv \sum_{m=1}^{p-1} \left(\frac{\lambda(m, rt)}{rt} + \lambda(m, 1) \right), \quad (11a)$$

modulo p.

A simpler property of the $\lambda(m, r)$ is given by

$$\lambda(m, m - r) = m - \lambda(m, r), \quad (12)$$

in view of (4).

Prof. Carlitz, after looking over our paper C, pointed out to me that the formula (6a) of that paper could be derived immediately from the known formula

$$(b - 1)^n - b_n = n(-1)^n, \quad (13)$$

by putting $n = p - 1$, and using

$$\binom{p-1}{i} \equiv (-1)^i (\text{mod } p). \quad (13a)$$

This proof is even simpler than the one referred to in the second footnote of C. He also showed that from (13) of the present paper, after setting $n = p$, we may obtain a result which is the formula (20) of C for $m = 1$, $r = 1$. By an extension of his method it is possible to prove a generalization of this formula (20) as we shall later show. In the original proof of (20) of C a number of known theorems concerning Wilson's quotient, Fermat's quotient and the Bernoulli numbers were employed, whereas in the present proof we shall see, aside from Fermat's theorem, that the only known results employed are (21a) together with

$$(b + t)^n - b_n = n(1^{n-1} + 2^{n-1} + \ldots + (t - 1)^{n-1} \quad (14)$$

and[2]

$$\frac{pb_{p-1} + 1}{p} \equiv W + 1 (\text{mod } p). \quad (15)$$

After (15) of C or (1) of the present paper is thereby proved, a number of the theorems employed in the first proof will follow as corollaries. The argument depends on finding two expressions for

$$\frac{(mb + k)^p}{p},$$

modulo p, where, as in the definition of $b_1, b_2, \ldots b_n$, we expand $(mb + k)^p$ by the binomial theorem and set b_h for b^h in the result. We have, if k can be expressed in the form $mt + ps$, with s and t integers,

$$\begin{aligned}(mb + k)^p &= ((mb + mt) + ps)^p \\ &= (mb + mt)^p + p^2s(mb + mt)^{p-1} + \left(\frac{p}{2}\right) p^2s^2 \\ &\qquad (mb + mt)^{p-2} + \ldots. \quad (16)\end{aligned}$$

Now, in this development, every term after the second is divisible by p^2 since p does not occur in the denominator of any Bernoulli number found in the expansion of $(mb + mt)^n$ for $n < p - 1$. This follows from the definition of the b's. Hence we may write

$$(mb + k)^p \equiv (mb + mt)^p + p^2s(mb + mt)^{p-1} \quad (17)$$

modulo p^2. Now in the same way

$$p^2s(mb + mt)^{p-1} \equiv p^2sm^{p-1}b_{p-1} \equiv pb_{p-1} \cdot ps \equiv -ps,$$

modulo p^2, using $pb_{p-1} \equiv -1 \pmod{p}$, which follows from (15). Also

$$\begin{aligned}(mb + mt)^p &= m^p(b + t)^p, \\ &= m^pb^p + pm^p(1^{p-1} + 2^{p-1} + \ldots + (t - 1)^{p-1}), \\ &\equiv pm\left(t - 1 - \left[\frac{t}{p}\right]\right), \quad (18)\end{aligned}$$

modulo p^2, for $t > 0$, using $b_p = 0$ for p odd. Hence (17) gives, $[x]$ being the greatest integer in x, $(mb + k)^p \equiv -ps + pmt - pm - pm\left[\frac{t}{p}\right]$, modulo p^2, and since $mt = k - ps$

$$\frac{(mb + k)^p}{p} \equiv -s - m + k - m\left[\frac{t}{p}\right] \pmod{p}. \quad (19)$$

On the other hand we may write by ordinary expansion

$$(mb + k)^p = p(mb)^{p-1} + k^p + pM, \quad (20)$$

where

$$M = \sum_{n=1}^{p-2} \binom{p}{n} m^n b_n k^{p-n}.$$

Now

$$\frac{pm^{p-1}b_{p-1}k + k^p}{p} = \frac{k(pb_{p-1}(m^{p-1} - 1) + pb_{p-1} + k^{p-1} - 1)}{p},$$

$$\equiv k\left(pb_{p-1}q(m) + q(k) + \frac{pb_{p-1} + 1}{p}\right),$$

$$\equiv k(-q(m) + W + 1 + q(k)), \tag{21}$$

modulo p, by (15) and noting that from (15) we obtain $pb_{p-1} \equiv -1 \pmod{p}$, with $(m, p) = 1$, $(k, p) = 1$.

We selected, in the above argument, k such that $k = mt + ps$ with $t > 0$. We may select $t > 0$ such that $mt \equiv k \pmod{p}$, since $(m, p) = 1$. Hence k may be any integer such that $(k, p) = 1$. Hence from (19), (20) and (21) we have, noting that

$$\binom{p}{n} \equiv (-1)^{n+1}\frac{p}{n} \pmod{p^2};\ q\left(\frac{m}{k}\right) - W + \sum_{n=1}^{p-2}\binom{m}{k}^n \frac{b_n}{n} \equiv \frac{1}{k}\left(s + \left[\frac{t}{p}\right]\right). \tag{22}$$

modulo p, where $(m, p) = 1$, $(k, p) = 1$, $k = ps + mt$, s, m and t are integers, $t > 0$, $q(x)$, W and b defined as in connection with (1).

The relation (22) is an extension of (1). For if we introduce the conditions stated in connection with (1), namely, $m > 0$, $0 < r < p$, and set $k = r$ in (22) then in $r = ps + mt$, $s < 0$ since $t > 0$, so if s_1 is now the original s in (4) we have, since $0 < s_1 < m$; $s_1 = m + s$ which reduces (22) to (1). By the use of (22) such relations as (11a) may be generalized.

We close with a remark of an elementary nature. In the relation (22) we reduced binomial coefficients modulo p. In a similar way we obtain, if

$$\binom{0}{0} = 1, \binom{a}{b} = 0 \text{ for } b > a,$$

$$(-1)^a\binom{p-1-a}{b} \equiv (-1)^b\binom{p-1-b}{a},$$

modulo p, for $0 \leqq b < p$; $0 \leqq a < p$. This is a "law of reciprocity" which I have not observed in the literature.

[1] These PROCEEDINGS, **34**, 103–110 (1948). This paper will be referred to as C.

[2] For other properties of Wilson's quotient cf. Emma Lehmer, *Ann. Math.*, **39**, 352 (1938).

[3] One proof of (15) is given in Vandiver, *Duke Math. Jour.*, **8**, 578 (1941), as a special case of a more general theorem. Other proofs are known.

SECOND ORDER DIFFERENTIAL EQUATIONS OF RANK 2 IN PHYSICS

BY EDWARD KASNER AND DON MITTLEMAN

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated March 30, 1949

1. Introduction.—All one-parameter families of plane curves are equivalent under point transformation. Thus, under the group of point transformations in the plane, there is no invariant of a one-parameter family of curves. However, under this same group, two-parameter families of curves (differential equations of second order) do possess invariants. Kasner[1] has determined the fundamental absolute invariants of differential elements of second order at a point. In the course of this investigation it was found expedient to introduce the homogeneous coördinates

$$z_1 : z_2 : z_3 : z_4 : z_5 = 1 : y' : y'^2 : y'^3 : y''.$$

This suggested the classification of differential equations of the second order, algebraic in y' and y'', according to the degree of the corresponding equation in the z coördinates. To an equation of first degree

$$A_1 z_1 + A_2 z_2 + A_3 z_3 + A_4 z_4 + A_5 z_5 = 0$$

there corresponds a differential equation of the form

$$Ey'' = Ay'^3 + By'^2 + Cy' + D,$$

where the coefficients are functions of x, y. This equation of rank 1 has been extensively studied by Lie, R. Liouville, Tresse and Kasner. It includes, among others, the differential equations for the geodesics of a surface, velocity families, natural families, isogonal families and Γ families.

The *rank* of a differential equation of second order, algebraic in y' and y'', is defined as the degree of the corresponding z equation.

The rank of a differential equation is invariant under an arbitrary point transformation.

The differential equations of rank 2 are:

$$A_0 y''^2 + (B_0 + B_1 y' + B_2 y'^2 + B_3 y'^3) y'' + (C_0 + C_1 y' + C_2 y'^2 + C_3 y'^3 + C_4 y'^4 + C_5 y'^5 + C_6 y'^6) = 0,$$

where the coefficients are arbitrary functions of x, y.

2. Central Loci.—Through a fixed point there pass integral curves of a given differential equation. The locus of the centers of the circles of curvature, constructed at a fixed point, to the integral curves passing through the point, is termed the *central locus*. With the fixed point as origin, the central locus, for equations of rank 1, is the cubic

$$E(x_0^2 + y_0^2) = -Ax_0^3 + Bx_0^2y_0 - Cx_0y_0^2 + Dy_0^3.$$

Such a cubic is characterized by the fact that the origin (the fixed point) is a double point and the two tangents are the minimal lines.

The only non-trivial degeneracy of this cubic occurs when $A = C$ and $B = D$, in which case the central locus is a straight line. A two-parameter family of curves for which, at each point, the central locus is a straight line is a *velocity family (of rank 1)*.

For equations of rank 2, the central locus of the $2\infty^1$ integral curves through a fixed point is the special sextic:

$$A_0[x_0^2 + y_0^2]^2 + (B_0y_0^3 - B_1x_0y_0^2 + B_2x_0^2y_0 - B_3x_0^3)[x_0^2 + y_0^2] + (C_0y_0^6 - C_1x_0y_0^5 + C_2x_0^2y_0^4 - C_3x_0^3y_0^3 + C_4x_0^4y_0^2 - C_5x_0^5y_0 + C_6x_0^6) = 0.$$

This sextic is characterized by the fact that the origin (the fixed point) is a quadruple point, the tangent lines are the two minimal lines, each counted twice, and they intersect the sextic only at the origin.

There is non-trivial degeneracy of the sextic when $C_3 = C_1 + C_5$ and $C_0 + C_4 = C_2 + C_6$. In this case the central locus is the special quartic characterized by the fact that the origin is a double point and the two tangents are the minimal lines. Further degeneracy occurs if $B_0 = B_2$, $B_1 = B_3$, $C_1 = C_5$, and $C_0 + 2C_6 = C_4$. In this case, the central locus is the conic:

$$A_0 + B_0y_0 - B_1x_0 + C_6x_0^2 - C_5x_0y_0 + C_0y_0^2 = 0. \qquad (1)$$

An important subset of differential equations of rank 2 for which the central loci are conics is that for which, at each point, the focus of the conic is at the point. In this case, the central loci are *focal conics* and the integral curves are termed a *velocity family of rank 2*. The conditions that the focus of the conic (1) be at the origin are $B_0B_1 = 2A_0C_5$ and $B_0^2 - B_1^2 = 4A_0(C_0 - C_6)$.

The differential equations of the $2\infty^2$ curves of a velocity family of rank 2 are

$$y'' = (\mathfrak{A} + \mathfrak{B}y')(1 + y'^2) \pm \mathfrak{C}(1 + y'^2)^{3/2},$$

where $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ are functions of x, y.

3. Differential equations of rank 2 arise from many diverse problems in geometry and physics. It is proposed to indicate some of these.

4. Scale Curves.[2]—Let $\mathbf{T}$ be a non-conformal mapping of the surface $\sum$, whose first fundamental form is

$$dS^2 = E(x, y)dx^2 + 2F(x, y)dx\,dy + G(x, y)dy^2,$$

upon the plane π, for which $ds^2 = dx^2 + dy^2$. The curves, on Σ or π, along which $\sigma^2 = (dS/ds)^2$ is constant, are scale curves. The totality

of these scale curves, for a specific mapping **T**, satisfies a differential equation of rank 2.

5. The two-parameter family of curves whose finite equation is either of the two distinct types

$$f_1a^2 + f_2a + f_3b + f_4 = 0$$
$$f_1ab + f_2a + f_3b + f_4 = 0,$$

where the f's are functions of x, y and a, b are the parameters, satisfies a differential equation of rank 2.

6. Any set of ∞^2 curves which is equivalent under a point transformation to the ∞^2 circles of unit radius satisfies a differential equation of rank 2.

Any set of ∞^2 curves which is conformally equivalent to the circles of unit radius is a velocity family of rank 2.

7. If $v(xy)$ is a scalar point function, $\mathbf{A}$ a vector point function, $d\mathbf{s}$ the vector arc length, ds the scalar arc length, then the system of curves which are extremals of the variational problem

$$\delta \int (v ds + \mathbf{A} \cdot d\mathbf{s}) = 0$$

is a velocity family of rank 2.

8. *Electric Particle Trajectories.*[3]—An electrically charged particle moves in a static electromagnetic field. If the motion is planar, the totality of trajectories, corresponding to all possible initial conditions, constitutes a three-parameter family of curves. If the electrostatic field is derivable from a potential, these curves may be grouped into ∞^1 families of ∞^2 trajectories each. Each family of ∞^2 trajectories is a velocity family of rank 2.

9. *Motion in a Rotating Plane and the Restricted Problem of Three Bodies in the Plane.*[4]—Let there be defined in the plane two rectangular Cartesian coördinate systems which have a common origin but one of which rotates, with respect to the other, with constant angular velocity ω. In the rotating coördinate system, the equations of motion for a particle of unit mass are

$$\ddot{x} - 2\omega\dot{y} = F \qquad (2)$$
$$\ddot{y} + 2\omega\dot{x} = G,$$

where F and G may be thought of as representing the components of a force in the x and y directions, respectively. In general, F and G depend on x, y, $\dot{x}$, $\dot{y}$ and t. The field will be called *positional* when F and G depend only on x, y. The totality of trajectories of a particle moving in such a positional field constitutes a three-parameter family of curves. If the field is derivable from a potential $-H(xy)$, then the equations (2) admit a

first integral

$$\dot{x}^2 + \dot{y}^2 + 2H = 2E,$$

where E is a constant of integration. In this case, the set of ∞^3 trajectories is composed of ∞^1 families of ∞^2 trajectories each, each family being associated with a specific value of E. The family of trajectories so determined is a velocity family of rank 2.

The classic example of the *restricted problem of three bodies* in the plane lies within this category.

10. The reason for calling a two-parameter family of curves whose central loci are focal conics a velocity family of rank 2 is made apparent by the following considerations.

In a plane, the intrinsic equations of motion of a particle of unit mass are

$$\frac{d^2s}{dt^2} = T \tag{3}$$

$$\frac{v^2}{r} = N, \tag{4}$$

where s is the arc length, r the radius of curvature of the trajectory, v the velocity, t the time, and T and N the tangential and normal components of the force. A trajectory is uniquely determined by the initial position, initial direction and initial speed. For the actual trajectory, the speed varies from point to point. If, in equation (4), the square of the speed, v^2, is replaced by a constant, say c^{-1}, the resulting equation is

$$\gamma = cN, \tag{5}$$

where, of course, γ is curvature. The curves which satisfy this differential equation are not, in general, trajectories. When the normal component of the force, N, is of the form

$$N = \varphi \cos \mu + \psi \sin \mu,$$

where μ is the inclination of the normal to the curve, φ and ψ point functions, and the constant c equals unity, the two-parameter family of curves so defined by (5) has been called by Kasner[5] a velocity family (of rank 1).

When N is of the form

$$N = \varphi \cos \mu + \psi \sin \mu + \chi,$$

where χ is a point function and the other symbols retain their meaning, the two-parameter family of curves defined by (5) is a velocity family of rank 2.

[1] Kasner, E., "The Geometry of Differential Elements of the Second Order with Respect to the Group of All Point Transformations," *Am. J. Math.*, **28**, (3), 203–213 (1907).

[2] Kasner, E., and DeCicco, J., "Scale Curves in General Cartography," PROC. NAT. ACAD. SCI. **30**, 211–215 (1944). "General Theory of Scale Curves," *Am. J. Math.*, **68**, No. 1 (1946).

[3] Lorentz, H. A., *Enc. d. math. Wiss.*, **14**, 156. Pauli, W., Jr., *Ibid.*, **19**, 634.

[4] Moulton, F. R., *An Introduction to Celestial Mechanics*, Revised, The Macmillan Company, New York, 1914, pp. 277–281.

[5] Kasner, E., "Differential-Geometric Aspects of Dynamics," *Am. Math. Soc. Colloquium Publication*, Vol. 3 (1912). A series of papers in the *Trans. Am. Math. Soc.*, Vols. 7–11 (1906–1911). Kasner, E., and DeCicco, J., "The Geometry of Velocity Systems," *Bull. Am. Math. Soc.*, **49** (4) (1943).

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The PROCEEDINGS will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be comprehensive in giving the relation of the paper to previous publications of the author or of others and in exhibiting, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work and long tables of data should be avoided, but authors should be precise in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though more elaborate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

REFERENCES to literature, numbered consecutively, will be placed at the end of the article and short footnotes should be avoided. It is suggested that references to periodicals be furnished in some detail and in general in accordance with the standard adopted for the Subject Catalogue of the International Catalog of Scientific Literature, viz., name of author, with initials following (ordinarily omitting title of paper), abbreviated name of Journal, with place of publication, series (if any), volume, inclusive pages, year. For example: Montgomery, T. H., *J. Morph., Boston*, 22, 731–815 (1911); or, Wheeler, W. M., *Königsberg, Schr. physik. Ges.*, 25, 1–142 (1914).

MANUSCRIPTS may be sent to the Managing Editor, Edwin Bidwell Wilson, Harvard School of Public Health, 695 Huntington Ave., Boston 15, Mass., only by members of the ACADEMY or by chairmen of the divisions of the NATIONAL RESEARCH COUNCIL who will assume responsibility for the propriety and scientific standards of the paper.

PROOF will not ordinarily be sent; if an author asks for proof, it will be sent with the understanding that charges for his corrections shall be billed to him. Authors are therefore requested to make final revisions on the typewritten manuscripts. The editors cannot undertake to do more than correct obvious minor errors.

REPRINTS should be ordered at the time of submission of manuscript. They will be furnished to authors at cost, approximately as follows:

SCHEDULE OF RATES FOR REPRINTS, CARRIAGE CHARGES EXTRA

No. Copies	50	100	150	200	250
4 pp.	$2.25	$2.75	$4.10	$4.85	$ 5.15
8 pp.	3.55	7.45	8.40	9.20	10.55
Covers	2.40	3.15	3.80	4.50	5.20

CONTENTS